

**ECONOMIC EVALUATION OF THE EFFECTS OF  
EARLY CHILDHOOD INTERVENTION PROGRAMS  
ON ADOLESCENT OUTCOMES**

Matthew Manning  
BCom, MA (Hons)

Key Centre for Ethics, Law, Justice and Governance,  
Faculty of Arts,  
Griffith University, Brisbane, Australia

Submitted in fulfillment of the requirements of the degree of Doctor of Philosophy

April, 2008

## **ABSTRACT**

Longitudinal research reveals that well designed and carefully implemented early-in-life intervention programs can produce positive short- and long-term benefits for at-risk children and their families. Benefits include gains in intellectual and academic achievement scores, improvements in educational outcomes, reductions in behavioural problems and delinquency, and improved family wellbeing. Convincing as this research is, not all programs produce benefits to the individual and his/her family; nor may they each produce net savings with respect to the costs of the criminal justice or education systems, or be cost-effective in terms of improved quality of life, feelings of safety, and enhanced wellbeing and happiness.

Economists have applied cost-benefit methodologies to measure the short- and long-term economic impacts of particular early childhood intervention programs. However, most economic analyses of such programs have focused on the governmental perspective (e.g. savings to the criminal justice system) with little emphasis placed on a more holistic, individual-level approach that measures benefits across multiple domains of life. Moreover, policy-makers have had access to a limited set of economic tools to assist them in making well-informed policy choices. This is particularly salient given the multiplicity of criteria that must be taken into account when making judgments that potentially have large effects on individuals and their families.

There are two prominent methodological deficiencies in the developmental prevention literature. The first deficiency is the limited array of methodological tools available to assist when making choices on resource allocation and engaging in a

structured decision-making process with respect to alternative policy options for early childhood interventions. The second deficiency is the absence of a rigorous tool for measuring the economic impact of early childhood interventions on salient aspects of non-health related quality of life throughout an individual's life, such as educational success, cognitive development, and social-emotional development. These aspects of quality of life are the primary focus of early intervention and developmental prevention programs.

In this thesis methods are developed to address both deficiencies. In the first study, a meta-analysis is conducted of the longitudinal research on the impact of early childhood interventions on the adolescent life phase. This study includes a detailed analysis of the psychometric properties of outcome measures relating to individuals' cognitive, social, and emotional development. Adolescence was selected because of the richness of follow-up data available for this life phase compared to other life phases (e.g. adulthood 28+ years), and because governments invest heavily in policies designed to combat adolescent problems such as delinquency, juvenile crime, drug abuse and conflict at school. A further practical reason for the adolescent focus is the complexity of gathering utility values for all life phases.

The second study adapts the analytical hierarchy process to develop a method for making complex multi-criteria decisions with respect to policy options for early childhood interventions. This procedure permits 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. Additionally, the second study

provides 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.

The meta-analysis highlighted the effects that early childhood interventions have on seven outcome domains during adolescence. Results demonstrated that early childhood intervention programs had the largest effect on educational success during adolescence followed (in order) by social deviance, social participation, cognitive development, criminal justice outcomes, family wellbeing, and social-emotional development. The analysis also revealed that programs that incorporated a structured preschool or centre-based educational component yielded positive effects on the outcome domains educational success and cognitive development throughout the adolescent life phase. Programs with a follow-through component into the early primary school years (e.g. preschool to Grade 3) also displayed strong effects on educational success and cognitive development in adolescence. Additionally, programs whose duration was longer than three years revealed larger sample means than programs that were longer than one year but shorter than three years. Program intensity was also found to be an important moderator of success. In combination, length and intensity of programs were important influences on the domains educational success, cognitive development, and deviancy.

In the second study, a survey of four stakeholder groups (policy people, people working in schools, a community agencies group, and an academic group) provided insights into priority rankings of alternative early childhood intervention programs

and relative utility values. It was found that family wellbeing was the highest priority with respect to its perceived contribution to non health-related quality of life during adolescence, with the child's social-emotional development the second highest priority.

When potential levels of program success were compared (from very high (VH) through H (high) to small (S) or no effect (N)), it was found that the larger the gap between levels (e.g., VH and S), the higher the preference score. Conversely, the smaller the gap between the levels of success (e.g. S and N) the smaller the preference score. Preference scores were not linear, with little discrimination between large effects (e.g. VH with H). This finding was consistent across all outcome domains. Respondents did not consider any level of success to be absolutely more important than any other, suggesting, in part, that they believed that some effect was better than no effect.

Study 2 revealed that a structured preschool program was considered the highest priority with respect to contributing to a strong effect on all outcomes during the adolescent years. This was followed in order of priority by family support services, parent education, centre-based childcare/ developmental day care, and home visitation. Using a hypothetical example, Study 2 also showed how relative utility values may be utilised to reveal the economic benefits (cost-utility) of early-in-life intervention programs on non health-related quality of life outcomes in the adolescent years.

Both studies had limitations that should be able to be overcome in future analyses. For example, it would have been beneficial in the meta-analysis to know how children from culturally and linguistically diverse groups differed in their responses to early-in-life intervention programs, and to have data on the effectiveness of early childhood intervention programs in different geographical locations. With respect to the second study (the analytical hierarchy process), a more detailed hierarchy that incorporated the most relevant indicators associated with the seven outcome domains (e.g. rates of special education, school graduation, or school drop-out) would have provided a more refined set of priority weights (or relative utilities) for adolescent outcomes. Further, incorporating elements such as length and intensity of program, and the use of follow-up, multi-component, or multi-contextual programs into the hierarchy would have been beneficial. These refinements were not possible in this study due to the survey procedures required to identify all the associated pairwise comparisons, which would have been extremely expensive and time consuming to conduct. Consequently, it is proposed that a series of separate follow-up studies be conducted using this methodology to identify relative utility values incorporating these additional elements. To overcome other limitations such as the relatively small number of participants surveyed under each category (e.g. policy-maker group, academic group etc.), a sensitivity analysis was conducted to measure the responsiveness of the results to changes in the relative importance of outcome objectives. This showed that results were very stable with respect to significant changes.

In summary, this thesis makes four significant contributions to the literature. First, it provides a meta-analytic overview of the outcomes associated with early

childhood intervention during the adolescent life phase. Outcome domains have been extended beyond those adopted in previous research to include educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes and family wellbeing and their respective indicators. Secondly, the analytical hierarchy process has been adapted to provide a systematic method for analysing policy decisions with respect to choices between early childhood intervention programs that potentially yield positive outcomes associated with enhanced quality of life during adolescence. Further, the adapted method provides decision-makers with a policy development mechanism whereby complex multiple criteria problems can be solved in a systematic way. The third original contribution of the thesis is a method for capturing individual relative utility values with respect to non health-related quality of life. Finally, a method for incorporating relative utility values into economic analyses has been outlined. The method provides options for measuring the cost-utility of early childhood intervention programs on the basis of a common metric outcome or set of outcomes that values qualitative improvements in an individual's life.

## DECLARATION OF ORIGINALITY

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

---

Signature

Matthew Manning

---

Date



## TABLE OF CONTENTS

ABSTRACT.....	I
DECLARATION OF ORIGINALITY .....	VII
LIST OF FIGURES .....	XVII
LIST OF TABLES .....	XXIV
ACKNOWLEDGEMENTS .....	XXVIII
CHAPTER 1: INTRODUCTION .....	1
1.1: Methodological gaps in the literature .....	1
1.2: Objectives and significance of this thesis .....	8
1.2.1: Objectives of the thesis .....	8
1.2.2: Significance of this study .....	9
1.3: Why the adolescent life phase? .....	12
1.4: Overview of the steps to overcome the methodological gaps .....	13
1.4.1: Section A (Chapters 2-4) .....	15
1.4.2: Section B (Chapters 5-6) .....	16
1.4.3: Section C (Chapters 7-10) .....	17
CHAPTER 2: AN ECONOMIC MODEL TO AID IN POLICY DECISION- MAKING AND MEASURING THE COST-UTILITY OF EARLY CHILDHOOD INTERVENTION .....	23
2.1: Introduction .....	23
2.2: Current methods for economic analysis of intervention programs .....	24
2.3: A methodological gap: The costs of crime .....	30
2.3.1: Individual benefits versus crimes averted? .....	31
2.3.2: Society or government? .....	36

2.3.3: The crime rate and its social consequences or the criminal event and consequences to the victim?.....	36
2.3.4: A second methodological gap: Making complex funding decisions under uncertainty .....	38
2.4: Conclusion .....	39
CHAPTER 3: PREVENTION, EARLY INTERVENTION AND THE DEVELOPMENTAL PERSPECTIVE.....	41
3.1: Introduction.....	41
3.2: Prevention and early intervention .....	42
3.2.1: The importance of early childhood intervention.....	43
3.2.2: The impact of early childhood intervention on at-risk children and their families.....	46
3.3: The Developmental Pathways model .....	48
3.3.1: Defining developmental prevention.....	48
3.3.2: Life phases, points of change and transition points .....	50
3.4: Reviews of the literature on the effectiveness of preschool prevention programs	52
3.5: The transition from childhood to adulthood .....	55
3.5.1: Defining adolescence .....	56
3.5.2: Transition points during the adolescent life phase.....	57
3.5.3: Interrelated conceptual domains of risk and protective factors during adolescence .....	59
3.5.4: Adolescent outcome domains .....	60
3.6: Conclusion .....	65

CHAPTER 4: QUALITY OF LIFE AND THE MEASUREMENT OF OUTCOMES ASSOCIATED WITH EARLY CHILDHOOD INTERVENTION DURING ADOLESCENCE.....	66
4.1: Introduction.....	66
4.2: Quality of Life (QOL): Health and non-health related .....	66
4.2.1: Domains of Quality of Life.....	70
4.2.2: Non-health related quality of life (NHRQOL) .....	71
4.3: Measuring the domains of NHRQOL and mapping their interactions .....	74
4.4: Instruments used to operationalise health and non-health related QOL domains	79
4.5: Instrument to gauge improvements in NHRQOL during adolescence resulting from early childhood intervention programs .....	80
4.5.1: Specifying measurement goals .....	81
4.5.2: Item generation .....	83
4.5.3: Item reduction and refinement .....	85
4.5.4: Questionnaire formatting .....	86
4.4.5: Instrument testing .....	88
4.5.6: Reliability.....	89
4.5.7: Consistency.....	89
4.5.8: Sensitivity analysis .....	89
4.5: Conclusion .....	90
CHAPTER 5: AN ANALYSIS OF LONGITUDINAL RESEARCH ON EARLY CHILDHOOD INTERVENTION PROGRAMS .....	92
5.1: Introduction.....	92
5.1.1: Overview of the meta-analytic methodology.....	92
5.1.2: Benefits of meta-analysis.....	95

5.1.3: Criticisms of meta-analysis.....	95
5.1.4: Effect size protocols used in meta-analysis .....	97
5.1.5: Calculation of individual effect sizes and variances.....	97
5.1.6: Calculation of weighted mean effect sizes for each predictor variable .....	99
5.1.7: Calculation of 95% confidence intervals .....	100
5.1.8: Calculation of the median effect size.....	100
5.1.9: Assessing heterogeneity of effect sizes using the Q-statistic .....	101
5.1.10: Fail-safe n .....	103
5.1.11: Interpretation of Meta-Analytic Results .....	104
5.2: Previous meta-analyses of early intervention programs for children .....	104
5.3: Outline of particular meta-analytic methodology employed in this thesis .....	106
5.3.1: Aim of the analysis .....	106
5.3.2: Identifying outcomes relevant to early childhood intervention across the life course.....	107
5.3.3: Meta-analytic framework.....	108
5.3.4: Selection criteria for inclusion in meta-analysis.....	109
5.3.5: Unit of analysis for relevant intervention programs .....	112
5.3.6: Preschool programs analysed.....	119
5.3.7: Variables coded in meta-analysis .....	123
5.3.8: Operationalisation of outcome domains .....	124
5.3.9: Effect size protocols used in this meta-analysis .....	127
5.3.10: Research questions.....	128
5.4: Conclusion .....	129

CHAPTER 6: RESULTS: A META-ANALYSIS OF LONGITUDINAL RESEARCH ON THE ADOLESCENT OUTCOMES ASSOCIATED WITH EARLY CHILDHOOD INTERVENTION PROGRAMS .....	131
6.1: Introduction.....	131
6.2: Meta-analytic results - coding and descriptive statistics .....	132
6.3: Mean effect size and Q statistic results.....	141
6.3.1: Discussion - Weighted mean effect size .....	143
6.3.2: Ranges of effect size .....	144
6.3.3: Effect sizes for ES and CD during adolescence for programs with and without a preschool education component.....	146
6.3.4: Effect sizes for early childhood intervention programs with and without follow through.....	147
6.4: Program characteristics as moderators of outcome effect sizes .....	150
6.4.1: Program length.....	150
6.4.2: Program intensity .....	154
6.4.3: Number of program components .....	157
6.5: Fail-safe n .....	159
6.6: Limitations of our meta-analysis .....	160
6.7: Implications of our meta-analysis.....	161
6.8: Conclusion and the link to part 2 of the thesis.....	163
CHAPTER 7: ECONOMIC ANALYSIS: PREFERENCES AND UTILITIES .....	166
7.1: Valuation of alternate life courses .....	167
7.2: Cost-effectiveness analysis .....	170
7.3: Cost-utility (CU) analysis .....	173
7.4: Utility Theory .....	176

7.4.1: Individual utility functions.....	179
7.4.2: Social utility function.....	181
7.4.3: Pareto optimality.....	182
7.5: Methods of measuring preferences .....	184
7.5.1: Rating scales and their variants .....	187
7.6.2: Standard gamble (health related) .....	189
7.6.3: Time trade-off (health related).....	191
7.5.4: Multi-attribute utility .....	193
7.6: Choosing the best method.....	195
7.7: The analytical hierarchy process (AHP) and preferences.....	197
7.8: Conclusion .....	199
 CHAPTER 8: THE ANALYTIC HIERARCHY PROCESS (AHP): AN OVERVIEW .....	 201
8.1: Introduction.....	201
8.2: Organising complex decisions into a hierarchical framework.....	202
8.2.1: Decomposition of decision .....	203
8.2.2: Comparative judgements within the hierarchy .....	203
8.2.3: Synthesis of priorities .....	206
8.3: Benefits and criticisms of the AHP method.....	210
8.3.1: Criticism 1 - Axiomatic foundations of AHP .....	211
8.3.2: Criticism 2 - Ambiguity of the questions that the decision-maker must answer .....	213
8.3.3: Criticism 3 - The scale used to measure the intensity of preference .....	214
8.3.4: Criticism 4 - The principle of hierarchical composition and rank reversal	216
8.4: An illustrative example of Saaty's method.....	217

8.5: Conclusion .....	228
CHAPTER 9: DEVELOPING THE ANALYTIC HIERARCHY PROCESS FOR USE IN EVALUATION OF ALTERNATIVE EARLY-IN-LIFE INTERVENTION PROGRAMS.....	229
9.1: Introduction.....	229
9.2: Why should we have to make a choice? .....	229
9.3: Application of the Analytical Hierarchy Process (AHP) Method .....	231
9.3.1: Step 1 - Decision problem .....	232
9.3.2: Step 2 - Framework for selection and data collection .....	239
9.3.3: Step 3 - Scenario of associated outcomes .....	249
9.3.4: Step 4 - Setting up the decision hierarchy .....	249
9.3.5: Step 5 - Pair-wise comparisons.....	253
9.3.6: Step 6: Estimating relative weights .....	253
9.3.7: Step 7 - Calculating consistency to validate results.....	255
9.3.8: Step 8 - Synthesise judgements/ Calculate relative weights of alternatives to obtain overall priorities indicating program desirability .....	256
9.4: Conclusion .....	260
CHAPTER 10: RESULTS FROM APPLICATION OF THE ANALYTIC HIERARCHY PROCESS TO EVALUATION OF ALTERNATIVE EARLY INTERVENTION PROGRAMS .....	262
10.1: Introduction.....	262
10.2: Results of survey 1 .....	263
10.2.1: Results of survey 1 section 1 .....	263
10.2.2: Results of survey 1 section 2 .....	266

10. 3: Results of survey 2.....	287
10.4: Synthesis of surveys 1 and 2 highlighting priorities for five early childhood intervention programs .....	299
10.5: Sensitivity analysis .....	305
10.6: Discussion of results .....	331
10.7: Limitations of the study .....	337
10.8: Conclusion .....	339
CHAPTER 11: SUMMARY AND CONCLUSION .....	341
11.1 Introduction.....	341
11.2: Results of the meta-analysis .....	342
11.3: Results of the analytical hierarchy process.....	346
11.4: Directions for future research .....	353
11.4.1: Economic appraisal of intervention programs .....	353
11.4.2: Applying a cost-utility approach.....	358
11.4.3: Measuring utility values.....	360
11.4.4: Using preference values to measure the economic impact of early childhood intervention programs on qualitative improvements in non-health related quality of life. ....	364
11.4.5: Identifying outcomes relevant to early childhood intervention across the life course .....	365
11.4.6: Appropriate methods for measuring preference/utility values of non health-related outcomes .....	366
11.4.7: Measuring the cost-utility of early childhood intervention programs – A hypothetical example .....	367
11.5: Limitations .....	374



11.6: Concluding remarks .....	378
APPENDICES .....	382
REFERENCE LIST .....	383

## LIST OF FIGURES

Figure 1.1: Schematic overview of thesis .....	14
Figure 3.1: A conceptual framework for adolescent risk behaviour: risk and protective factors, risk behaviours, and risk outcomes .....	61
Figure 4.1: Levels of quality of life .....	71
Figure 4.2: Measuring quality of life using the Spilker model .....	74
Figure 4.3: Factors which may influence the various quality of life domains .....	77
Figure 4.4: Domains and individual components in the current instrument .....	84
Figure 6.1: Weighted average effect sizes (d.) corrected for sample size for seven adolescent outcomes .....	141
Figure 6.2: Average weighted effect sizes for the outcome educational success (ES) for programs with and without a follow through component .....	148
Figure 6.3: Average weighted effect sizes for the outcome cognitive development (CD) for programs with and without a follow through component .....	149
Figure 6.4: Average weighted effect sizes for the outcome educational success for program duration <1 year >3 years and <3 years .....	151
Figure 6.5: Average weighted effect sizes for the outcome educational success (ES) for programs whose intensity was <500 sessions and >500 sessions .....	154
Figure 6.6: Average weighted effect sizes for number of program components .....	158
Figure 7.1: An illustration of Pareto efficiency .....	183
Figure 7.2: The visual analogue scale .....	188
Figure 7.3: Standard gamble for chronic health state preferred to death .....	189
Figure 7.4: Standard gamble for temporary health state .....	190
Figure 7.5: Time trade-off for a chronic health state preferred to death .....	192
Figure 7.6: Time trade-off for a temporary state .....	193

Figure 8.1: An example of a hierarchy .....	204
Figure 8.2: Advantages of the analytical hierarchy process (AHP) method.....	212
Figure 8.3: An illustrative hierarchy for determining preferences .....	218
Figure 8.4: Matrix comparing six outcome domains .....	219
Figure 8.5: Normalised columns of the matrix .....	220
Figure 8.6: Obtaining the vector of priorities .....	220
Figure 8.7: Determining preferences among the intensities for the outcome educational success (ES).....	221
Figure 8.8: Determining preferences among the intensities for the outcome cognitive development (CD).....	221
Figure 8.9: Determining preferences among the intensities for the outcome social- emotional development (SED).....	221
Figure 8.10: Determining preferences among the intensities for the outcome deviancy (D) .....	222
Figure 8.11: Determining preferences among the intensities for the outcome social participation (SP) .....	222
Figure 8.12: Determining preferences among the intensities for the outcome criminal justice outcomes (CJO).....	222
Figure 8.13: Matrix comparing intensity levels with respect to each outcome .....	223
Figure 8.14: Vectors of priority for the six outcome intensities .....	223
Figure 8.15: Normalised vector of priorities with respect to the most desired attribute- intensity outcome .....	224
Figure 8.16: Normalising the vector .....	224
Figure 8.17: Comparing options X, Y, and Z for desired attribute intensity H-ES ...	225
Figure 8.18: Comparing options X, Y, and Z for desired attribute intensity H-CD ..	225

Figure 8.19: Comparing options X, Y, and Z for desired attribute intensity L-SED.	225
Figure 8.20: Comparing options X, Y, and Z for desired attribute intensity H-D.....	225
Figure 8.21: Comparing options X, Y, and Z for desired attribute intensity M-SP...	226
Figure 8.22: Comparing options X, Y, and Z for desired attribute intensity H-CJO.	226
Figure 8.23: Overall option attributes perception - Matrix comparing attribute-intensity levels with respect to each alternative (X, Y, Z).....	226
Figure 8.24: Weighted overall option attribute perception.....	227
Figure 8.25: Vector of priorities for program alternatives based on pair-wise comparisons of the all levels of the hierarchy .....	227
Figure 9.1: Steps in the application of the analytical hierarchy process (AHP) method .....	232
Figure 9.2: Disadvantaged areas in Brisbane (1996 Index of Relative Socio-Economic Disadvantage SEIFA) .....	246
Figure 9.3: Decision hierarchy for early childhood intervention program alternatives, outcome domains and level of effect .....	252
Figure 10.1: Priority percentage for outcomes during adolescence.....	265
Figure 10.2: Pair-wise comparison between outcome levels of success very high effect (VH) and no effect (N).....	275
Figure 10.3: Pair-wise comparison between outcome levels of success high effect (H) and no effect (N) .....	276
Figure 10.4: Pair-wise comparison between outcome levels of success medium effect (M) and no effect (N).....	277
Figure 10.5: Pair-wise comparison between outcome levels of success small effect (S) and no effect (N) .....	278

Figure 10.6: Pair-wise comparison between outcome levels of success very high effect (VH) and small effect (S).....	279
Figure 10.7: Pair-wise comparison between outcome levels of success high effect (H) and small effect (S) .....	280
Figure 10.8: Pair-wise comparison between outcome levels of success medium effect (M) and small effect (S).....	281
Figure 10.9: Pair-wise comparison between outcome levels of success very high effect (VH) and medium effect (M).....	282
Figure 10.10: Pair-wise comparison between outcome levels of success high effect (H) and medium effect (M).....	283
Figure 10.11: Pair-wise comparison between outcome levels of success high effect (VH) and medium effect (H).....	284
Figure 10.12: Percentage of priorities of five early childhood intervention program alternatives with respect to a very high effect on seven outcome domains .....	299
Figure 10.13: Synthesis of priorities .....	300
Figure 10.14: Normalising of the synthesised priorities.....	301
Figure 10.15: Summary of the hierarchy incorporating results from level 1.....	302
Figure 10.16: Overall performance graph for hierarchy .....	304
Figure 10.17: Sensitivity for early childhood intervention alternatives (Reducing the most preferred outcome (FW) by 10%).....	307
Figure 10.18: Overall performance graph for hierarchy with a 10 percent reduction in the outcome FW .....	308
Figure 10.19: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome FW and possible changes in program priority rankings .....	309

Figure 10.20: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome CJO and possible changes in program priority rankings .....	310
Figure 10.21: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome ES and possible changes in program priority rankings .....	311
Figure 10.22: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome CD and possible changes in program priority rankings .....	312
Figure 10.23: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome SED and possible changes in program priority rankings .....	313
Figure 10.24: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome D and possible changes in program priority rankings.	314
Figure 10.25: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome SP and possible changes in program priority rankings	315
Figure 10.26: Head to head graph of SPP vs. CBCC on seven outcome domains ('what if' scenario 1).....	316
Figure 10.27: Head to head graph of SPP vs. HV on seven outcome domains ('what if' scenario 1).....	317
Figure 10.28: Head to head graph of SPP vs. FSS on seven outcome domains ('what if' scenario 1).....	317
Figure 10.29: Head to head graph of SPP vs. PE on seven outcome domains ('what if' scenario 1).....	318

Figure 10.30: Sensitivity for outcome domains and early childhood intervention alternatives (Reducing the most preferred alternative (FW) by 20%).....	319
Figure 10.31: Overall performance graph for hierarchy with a 20 percent reduction in the outcome FW.....	320
Figure 10.32: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome FW and possible changes in program priority rankings .....	322
Figure 10.33: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome CJO and possible changes in program priority rankings .....	323
Figure 10.34: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome SP and possible changes in program priority rankings.....	324
Figure 10.35: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome D and possible changes in program priority rankings.....	325
Figure 10.36: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome SED and possible changes in program priority rankings .....	326
Figure 10.37: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome ES and possible changes in program priority rankings .....	327
Figure 10.38: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome CD and possible changes in program priority rankings .....	328
Figure 10.39: Head to head graph of SPP vs. CBCC on seven outcome domains ('what if' scenario 2).....	329

Figure 10.40: Head to head graph of SPP vs. HV on seven outcome domains (‘what if’ scenario 2) .....	330
Figure 10.41: Head to head graph of SPP vs. FSS on seven outcome domains (‘what if’ scenario 2) .....	330
Figure 10.42: Head to head graph of SPP vs. PE on seven outcome domains (‘what if’ scenario 2) .....	331
Figure 11.1: Components of an economic appraisal.....	354
Figure 11.2: Quality adjusted life years (QALY’s) gained from an intervention.....	356
Figure 11.3: Example of attributes for a non health-related classification system....	363



## LIST OF TABLES

Table 2.1: Approaches to economic analysis of crime prevention programs .....	29
Table 3.1: Adolescent developmental transitions: Stressors, risk factors and preventative strategies.....	58
Table 3.2: Adolescent outcome domains and their indicators .....	63
Table 3.3: Projects and studies included in the meta-analysis, from which outcomes domains and their indicators were developed - Adolescent years .....	64
Table 4.1: Domains and components of non-health related quality of life.....	72
Table 4.2: Example of possible life events .....	76
Table 4.3: Generic steps in the development and validation of a new non health- related quality of life instrument.....	81
Table 4.4: Instrument development and validation in quality of life measurement ....	82
Table 5.1: Early intervention programs included in the meta-analysis .....	114
Table 5.2: Outcome domains and operationalisations across studies included in the meta-analysis.....	126
Table 6.1: Codes for individual (primary) studies included in the meta-analysis. ....	133
Table 6.2: Program characteristics of included studies .....	134
Table 6.3: Characteristics of program participants .....	135
Table 6.4: Study Characteristics .....	135
Table 6.5: Effect sizes (d and d.), variance, inverse variance, CI Lower and Upper (95%), corrected for small sample size, for the seven outcome domains during the adolescent life phase. ....	137
Table 6.6: Results - Test of homogeneity for seven outcome domains.....	143
Table 6.7: Statistics for ranges of effect size during adolescence .....	145
Table 7.1: Methods of measuring preferences and utilities .....	186

Table 8.1: Saaty's Comparison scale .....	204
Table 8.2: Example of pairwise comparisons of parties (A, B, C and D) .....	205
Table 8.3: Average random index (R.I) for matrices of order 1-15.....	210
Table 9.1: Indicators of the outcomes educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes, and family wellbeing .....	239
Table 9.2: Snapshot of the Inala community taken from the 2001 Census .....	246
Table 9.4: Matrix comparing outcome domains/attributes .....	253
Table 9.5: Matrix comparing intensity levels with respect to the outcome/attribute educational success (ES).....	255
Table 9.6: Priorities of attributes/outcomes with respect to five levels of intensity..	257
Table 9.7: Weighted vectors of priority for all levels of intensity.....	257
Table 9.8: Normalised vector of desired attribute/outcome intensities .....	258
Table 9.9: Matrix comparing the five early childhood intervention program alternatives for desired attribute/outcome intensities .....	258
Table 9.10: Weighted vector of priorities for desired attribute/outcome intensities for early childhood intervention alternatives.....	259
Table 9.11: Overall weighted vector of priorities for the five early childhood intervention program options.....	259
Table 10.1: Pair-wise comparison of adolescent outcomes (average response).....	264
Table 10.2: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome educational success (ES) .....	268
Table 10.3: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome cognitive development (CD) .....	269

Table 10.4: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome social-emotional development (SED) .....	270
Table 10.5: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome deviancy (D) .....	271
Table 10.6: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome social participation (SP) .....	272
Table 10.7: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome criminal justice outcomes (CJO) ..	273
Table 10.8: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome family wellbeing (FW) .....	274
Table 10.9: Summary table of the results comparing outcome levels of success with respect to their perceived importance on seven outcomes during adolescence.	285
Table 10.10: Vector of priorities for most preferred outcome/level of success.....	286
Table 10.11: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-ES. ....	289
Table 10.12: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-CD. ....	290
Table 10.13: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-SED. ....	290
Table 10.14: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-D. ....	292
Table 10.15: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-SP. ....	293

Table 10.16: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-CJO.....	295
Table 10.17: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-FW.....	295
Table 10.18: Weighted priorities for the early childhood intervention program options with respect to outcome/intensities .....	297
Table 10.19: Overall priorities of five early childhood intervention program alternatives with respect to very high effect on seven outcome domains.....	298
Table 10.20: Hypothetical cost-effectiveness scenario: Comparing the cost/effectiveness of 3 levels of effect (very high effect, medium and small effect).....	334
Table 11.1: Resources Used in the Delivery of an Intervention .....	370
Table 11.2: Worksheet for Estimating Costs .....	373
Table 11.3: Costs, utility and cost-utility ratio for three preschool program alternatives .....	374

## ACKNOWLEDGEMENTS

I would like to thank my parents for instilling in me the drive to succeed and teaching me the value of education, but more importantly, the value of hard work. Moreover, I would like to acknowledge their indefatigable support throughout my life. I will always be thankful to my wife, Annie, for her ongoing support and encouragement, and for motivating me to overcome any obstacles. There is no more beautiful person in the world than my Annie! I would also like to thank my sisters Lisa and Gabrielle and their husbands Matthew and Jonathan and my uncle, Trevor Jenkins, for always `being there` for me. I am also blessed to have the most wonderful sons, Thomas and Edward, who are kind, sincere and gentle young men.

I have been positively influenced in so many ways by the academics in the School of Criminology and Criminal Justice, and the Department of Accounting, Economics and Finance, since I began my university studies at Griffith University. Further, I am forever grateful to my supervisors and close friends – Professor Christine Smith, and Professor Ross Homel – whose guidance and support has helped me to better myself both academically and personally. Their support has provided me and my family with a life that would not have been possible without their input.

I would also like to acknowledge the support from other academics and peers, most notably Professor Richard Wortley, Professor Kay Ferres, Associate Professor Anna Stewart, Dr Susan Dennison, Professors Paul and Lorraine Mazerolle, Dr Janet Ransley, Dr Melissa Bull, Dr Kate Freiberg, Mr Kerry Wimshurst, Associate Professor Tim Prenzler, Dr Rebecca Wickes, Associate Professor Jay Bandaralage, Ms Celia Dickson, Mrs Lorraine Coutts, Ms Rebecca Glover, Ms Marianne Mitchell, and Ms Sandra Hart.

Finally, I would like to recognize my true and closest friend Matthew Seeto, a friend I am very blessed to have, and Sharyn and Maurice Pearce who have supported me through tough times and continue to be friends I hold close to my heart.

## **CHAPTER 1: INTRODUCTION**

### **1.1: Methodological gaps in the literature**

The benefits of well-designed and carefully implemented early childhood interventions programs have been confirmed by longitudinal research (Homel, 2005). This is particularly salient given reports demonstrating deterioration over time for Australian children with respect to health outcomes (e.g. asthma, diabetes, obesity, intellectual disability, eating disorders, depression, attention deficit disorder, hyperactivity), child abuse, neglect, crime and behaviour problems (Freiberg, Homel, & Lamb, 2007; Stanley, Richardson, & Prior, 2005). An influential series of experiments have produced solid evidence that intervening (with the child and the family) early in the developmental pathway can successfully head-off future health, educational, behavioural and crime-related problems (Farrington & Welsh, 2002).

Evaluations of the short- and long-term effects of early childhood interventions including projects such as the Perry Preschool Project (Schweinhart, 2004), the Elmira Prenatal/Early Infancy Project (Eckenrode, Olds, Henderson, Kitzman, Luckey, Pettitt, Sidora, Morris, Powers, & Cole, 1998; Olds, 2002) and the Seattle Social Development Project (Hawkins, Catalano, Kosterman, Abbot, & Hill, 1999) have confirmed the positive effects that well designed early childhood interventions can have on at-risk children and their families. This is 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 outcomes, a reduction in behavioural problems and delinquency and improved family wellbeing (Brooks-Gunn, Fuligni, & Berlin,

2003, p.5-9). Moreover, a systematic delivery of basic services or resources 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). 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 behaviour, lower levels of substance abuse, and increases in income and workforce participation (Brooks-Gunn et al., 2003; Olds, 2002; Reynolds, Temple, Robertson, & Mann, 2001).

Improvements in child behaviour, learning and health, together with benefits to families and society, create a compelling argument for the continued refinement and ongoing implementation of early childhood intervention programs. The amalgamation of three seemingly distinct policy areas (health, welfare and crime prevention), together with the creative application of economic evaluation, has further enriched the development of thought in this field. The emergence of a small, but significant literature (Aos, Lieb, Mayfield, Miller, & Pennucci, 2004; Barnett, 2000; Belli, Bustreo, & Preker, 2005; Eckenrode et al., 1998; Farrell, Bowers, & Johnson, 2004; Homel, Freiberg, Lamb, Leech, Carr, Hampshire, Hay, Elias, Manning, Teague, & Batchelor, 2006; Karoly, Greenwood, Everingham, Hoube, Kilburn, Rydell, Sanders, & Chiesa, 1998; Manning, 2004; Manning, Homel, & Smith, 2006; Schweinhart, 2004; Schweinhart et al., 1993; Xiang & Schweinhart, 2002) quantifying the economic benefits of developmental interventions has further propelled the argument for continual development and application of early childhood interventions. Future savings to the criminal justice system, a reduction in costs borne

by victims of crime, a decline in welfare assistance claimants and smaller numbers of children requiring special education provide examples of the economic benefits revealed in these studies.

However, the claim that all early childhood developmental programs will produce similar effects, or produce significant cost savings, cannot be convincingly sustained given the myriad of program types (e.g. structured preschool programs, home visitation and parental education), types of individuals being served (e.g. children of disrupted families, foster children, homeless children, bereaved children, children of symbiotic and psychotic parents, and neglected children), and geographical locations, (“differences based on geographic location [urban and rural], such as availability of resources, access to resources, and familial contact, may impact how services should be delivered”(Darling, 2003, p.1)).

This thesis developed as a result of a seminal paper by Professor Daniel Nagin (Nagin, 2001), proposing several possibilities that would facilitate improvements in current methods (e.g. cost-savings, cost-effectiveness and cost-benefit) applied by economists to measure the economic impact of crime prevention programs on outcomes in both the short- and long-term. Nagin’s work focused on issues regarding the theoretical and practical obstacles that one faces when estimating the benefits of crime prevention programs. Rather than criticising the more technical matters of economic analysis (e.g. choosing the right discount rate), he focused on the broader conceptual issues concerning the overall structure of an analysis. In short, Nagin proposed the need to develop a methodology that measures benefits across multiple domains, at different times, yet at the individual level. Further, Nagin posited that the



current methods applied by economists in the area of crime prevention/developmental intervention do not, to date, account for the valuable impacts these programs provide; for example, increased public safety, or salient individual benefits, such as improved quality of life. I agree with Nagin, proposing that the natural unit of analysis in measuring the economic effectiveness of early childhood intervention should be on the individual. Thus, economists should strive to measure the qualitative improvements (e.g. social, emotional, and developmental) in a child and his/her family's quality of life as a result of early intervention. A more detailed discussion on Nagins work is provided in Chapter 2.

Further to Nagins work, the impetus to develop a methodology that measures benefits to the individual and the family as a result of early-in-life interventions stemmed from my work on the Pathways to Prevention project (Manning et al., 2006). One of several limitations of the economic analysis of the Pathways project was the inability to conduct a cost-utility analysis. Cost-utility analysis allows one to evaluate program alternatives according to a comparison of their costs and perceived utility (Boardman, Greenberg, Vining, & Weimer, 2001). Essentially, the term utility refers to an individual's perceived relative satisfaction derived directly from use of goods and services as well as from the qualitative improvements in quality of life that such use engenders (Torrance, 1986). The significance of cost-utility analysis is that outcomes may be single or multiple, can be general as opposed to program specific and may incorporate a notion of value to both the individual and society (Drummond, O'Brien, Stoddart, & Torrance, 1997). This form of economic analysis moves beyond methods currently applied in the area of crime prevention/early intervention, whereby a set of common metric outcomes may be identified. Hence, dissimilar programs may

be compared from an economic perspective, where once direct comparisons were impossible.

For the purposes of this thesis, focus is specifically placed on the salient effects that early childhood intervention programs may have on an individual's quality of life. Quality of life is defined as a broad set of life domains that represent the effect of a series of states which impact, either positively or negatively, on an individual's life trajectory (Spilker, 1996b). It is acknowledged that quality of life incorporates both health and non-health related domains. However, in this thesis, the focal point is primarily non-health related outcomes. Although health-related outcomes are important to include into any economic analysis, it was found (through the application of a meta-analysis) that our current knowledge of what they are, and our ability to measure them is currently limited. Consequently, attempting to incorporate health-related measures into an economic analysis is limited until empirical evidence provides data highlighting the impact of early-in-life interventions on health-related outcomes across the life course. Additionally, the meta-analysis revealed the importance researchers place on non health-related outcomes (e.g. educational success, cognitive development, deviancy) and the impact non health-related outcomes have on quality of life throughout the life course. Therefore, based on the limitations imposed on economists to measure health-related outcomes into their analyses, and the importance of measuring non health-related outcomes, this thesis concentrates on measuring the impact early childhood interventions have on non health-related outcomes; leaving health-related outcomes for future research.

Part of this thesis incorporates a meta-analysis of longitudinal research into the impact of early childhood/developmental interventions on individual and family non health-related outcomes. The meta-analysis provided a basis from which to identify the salient outcomes of these interventions and the strength of those outcomes, with this knowledge being used in another part of the thesis to assist individuals in developing relative utility values for their perceived preferences with respect to potential outcomes at different times in the life course. Outcomes associated with the adolescent life phase were chosen as the focus for this thesis. This life phase was selected given that no previous meta-analyses adequately captured the salient outcomes associated with this life phase for at-risk populations. Moreover, it has been observed that considerable amount of interest exists on the part of stakeholders on outcomes associated with the adolescent years, for example reductions in delinquency and crime related outcomes. A meta-analysis was necessary to identify not only the salient outcome domains and their indicators, but the strength and degree of effectiveness that early childhood interventions had on improvements in quality of life during adolescence. Specifically, this research defines quality of life as a set of non health-related domains (e.g. cognitive, social-emotional) explicitly related to the outcomes associated with early childhood intervention on adolescence. The domains represent the effect of a series of states (e.g. school achievement tests, social skills, self-esteem, behaviour), which impact, either positively or negatively, on the life course trajectory of an individual and his/her family.

As a result of a detailed analysis of the literature (Chapters 2 and 3) and my work with academics and policy-makers on the Pathways to Prevention project, two methodological gaps were identified: (1) the array of tools able to be accessed to

make well informed choices on resource allocation and structured decision-making with respect to alternative policy options for early childhood interventions, and (2) our inability to accurately measure the economic impact of early childhood interventions on various outcomes associated with non-health related quality of life (e.g. educational success, cognitive development, social-emotional development) throughout an individual's life-course (e.g. early childhood and adolescence) (Nagin, 2001).

Obstacles exist with respect to making complex decisions without a structured methodological decision-making procedure. A detailed scan of the literature and subsequent meetings with policy-makers revealed that policy decisions in the area of crime prevention/early interventions are made without the use of a systematic decision-making methodology. Moreover, although decisions are often made based on empirical research, Saaty (2000) reveals that complex decisions of this nature cannot be made adequately without a detailed framework for capturing all the salient elements of the decision/problem. Consequently, it is argued that current policy-making decisions are imperfect and could be improved through the development and use of a structured procedure.

Further, measuring the economic impact of early childhood interventions on the salient life outcomes associated with non health-related quality of life (e.g. improvements in educational success, cognitive development, and social-emotional development) across an individual's life course is difficult, if not impossible, using current methods. Therefore, a method needs to be developed or adapted that first, provides a structured methodological procedure for assisting those responsible for

making complex multi-criteria decisions (incorporating educational, cultural and social and economic factors) with respect to policy options for early childhood interventions. Secondly, identifies a common metric outcome/s across competing and often disparate program alternatives (e.g. early intervention programs such as home visitation and centre-based developmental day care) with the goal of measuring the economic impact (in terms of cost-utility) of early childhood interventions on non health-related outcomes during an individual's life course.

Nagin (2001) argues that the development of such a method has the potential to be much more than a rhetorical tool. That is, something to produce an effect rather than providing an answer. In particular, this form of economic analysis has the potential to make a significant impact on crime control, educational and social welfare policy.

## **1.2: Objectives and significance of this thesis**

### **1.2.1: Objectives of the thesis**

With overcoming the above-mentioned methodological gaps as its primary focus, this thesis aims:

- to highlight the importance of economic analysis in developing and evaluating early childhood intervention programs, and identify potential outcomes (social and economic) associated with early childhood interventions across the life course;
- to identify outcome domains associated with early childhood intervention and their indicators during the adolescent life phase;

- to establish the degree of effectiveness (in terms of effect size) of five forms of early childhood intervention (structured preschool programs, centre-based child care/developmental day care programs, home visitation, family support services and parental education) on seven outcome domains (educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes and family wellbeing) during the adolescent life phase;
- to develop a methodology for obtaining utility values for decisions regarding options for early childhood intervention programs and their contribution to increasing non-health related quality of life during the adolescent life phase;
- to provide a methodology for analysing complex multiple criteria problems with the aim of providing a systematic decision making tool to aid in the policy planning process with respect to choices of early childhood intervention programs and their perceived contribution to increasing non health-related quality of life during the adolescent life phase;
- to provide options for measuring the cost-utility of early childhood intervention programs incorporating a common metric that properly values qualitative improvements in an individual's quality of life; and,
- to highlight areas of research for future study.

### **1.2.2: Significance of this study**

This thesis is significant for four reasons:

1. It provides a meta-analytic overview of the outcomes associated with early childhood intervention during the adolescent life phase. This is particularly significant given there appears to be no published summary of results from longitudinal research on a broad range of non health-related outcomes related to at-risk children and their families during the adolescent life phase. Outcome domains have been extended beyond those adopted in research conducted by Nelson, Westhues, Laqurier, and MacLeod (2003) to include educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes and family wellbeing and their respective indicators. The methodology employed in this meta-analysis has been improved compared with previous analysts to ensure that outcome domains measured and their indicators are grouped according to methods employed by psychologists using psychometric tests of individuals of different ages.
2. It adapts the analytic hierarchy process (Saaty, 1980) method to provide a systematic way to analyse policy decisions with respect to choices between early childhood intervention programs that can potentially yield outcomes associated with quality of life in the adolescent life phase. This is particularly significant given there appears to be no published method available to assist policy-makers in making complex decisions regarding policy options for early childhood intervention programs. The analytical hierarchy process has been selected for adaptation as it provides decision-makers with a policy development mechanism whereby complex multiple criteria problems can be

solved in a systematic way. This is based on a belief that complex decisions of this nature cannot be made without a structured process that incorporates all elements of a decision into the decision framework. Moreover, we argue that the adapted methodology can be applied to decision-making in any area of criminological research to assist in making policy decisions regarding program options at any stage in the life course (e.g. early childhood, childhood, adulthood).

3. It develops a method for capturing individual perceived relative utility values with respect to choosing between early childhood intervention alternatives that potentially contribute to improvements in an individual's non health-related quality of life. This is significant since past economic analyses of crime prevention/early intervention programs have been unable to measure benefits across multiple domains, at different times, and with a focus on the level of the individual (Nagin, 2001). Moreover, Nagin highlights that current methods applied by economists in the area of crime prevention/developmental intervention do not account for many of the important impacts these programs provide; for example impacts related to improved quality of life.
4. It provides options for measuring the cost-utility of early childhood intervention programs ensuring that one identifies a common metric outcome or set of outcomes that values a qualitative improvement in an individual's quality of life. This is important as current methods employed by economists have not extended to incorporate cost-utility analysis. Therefore, a detailed example is provided to assist those who use the methodology adapted in this



thesis to apply the relative utility weights generated via the analytical hierarchy process to measure the utility of early-in-life intervention programs. These utility measures can then be compared with the costs of such programs using standard cost-utility methodology.

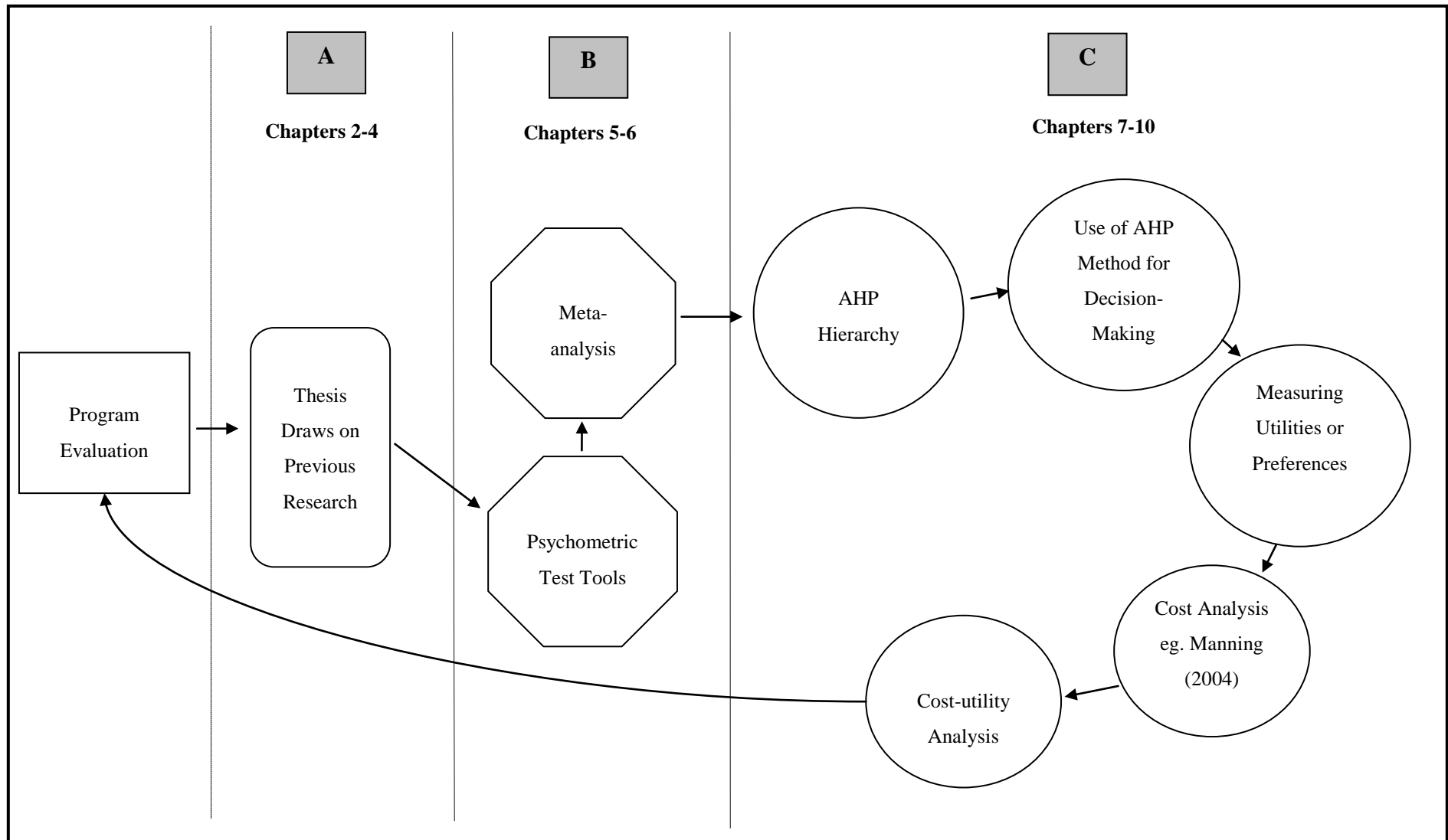
### **1.3: Why the adolescent life phase?**

The adolescent life phase was chosen as the focus of this thesis because of the richness of follow-up data available across this life phase compared to other life phases (e.g. adulthood 28+ years). Furthermore, previous analyses highlighting the results of early childhood interventions on non health-related outcomes during adolescence have not managed to adequately measure outcome domains beyond educational success and cognitive development. Moreover, indicators that make up the domains have often been chosen and categorised with little justification for their placement under those domains. In this thesis, all indicators are categorised (placed under the umbrella of a outcome domain) according to the results of a detailed analysis of the psychometric test tools used by psychologists to measure academic skills/educational success, behaviour at school, social-emotional development, family wellbeing, and self-reported deviance (Appendix A). Attention is focused on this life phase and its major transition points; particularly by government, whose interest often surrounds outcomes associated with this life phase. These associated outcomes include, delinquency, juvenile crime, drug abuse and conflict at school. Finally, research was limited to the adolescent years because of the complexity of gathering utility values for all life phases. Research suggests that limiting research to one life period will result in more consistency across responses (Gold, Siegel, Russell, &

Weinstein, 1996b). Consequently, it was decided that the methodology be tested first on a single life phase, with later research expanding to encompass other life phases.

#### **1.4: Overview of the steps to overcome the methodological gaps**

Given the multiple objectives of this thesis, it is organised into three sections. Figure 1.1 provides an overview of the three sections (A, B and C) of this thesis and identifies how all sections link to provide a method that could be used to improve program evaluation of early childhood interventions. Section 1.1 outlines the methodological gaps in the literature highlighting two proposed methodological improvements for early childhood intervention program evaluation: (1) a method for making well informed choices on resource allocation and structured decision-making with respect to alternative policy options, and (2) a method to measure the economic impact on various outcomes associated with non-health related quality of life (e.g. educational success, cognitive development, social-emotional development) throughout an individuals life-course (e.g. early childhood and adolescence). To address the above-mentioned methodological gaps, we begin at Section A, followed by Section B and finally Section C. Sections 1.4.1, 1.4.2, and 1.4.3 provide a description of Sections A, B and C respectively as well as a brief discussion highlighting the contents of each chapter in those Sections.

**Figure 1.1: Schematic overview of thesis**

### **1.4.1: Section A (Chapters 2-4)**

The thesis begins with a detailed review of the crime prevention, and early childhood intervention and developmental literatures. This review provides an overview of the gaps in the economic analysis of crime prevention and highlights the need to develop a structured decision-making tool for policy-makers when making complex multi-criteria decisions with respect to crime prevention policies. A review of the quality of life literature is provided highlighting the steps used to identify relevant quality of life outcome domains and their respective indicators. A method for measuring improvements in quality of life across an individual's life course is articulated.

Chapter 2 provides an overview of the crime prevention literature highlighting the need to develop an economic model that (a) provides a structured process for making complex decisions under uncertainty with respect to policy options involving early childhood intervention/crime prevention programs, and (b) is able to identify a set of common metric for outcomes that accurately value qualitative improvements in an individual's life from a cost-utility perspective.

Chapter 3 reviews the literature with respect to the developmental perspective, crime prevention, and early childhood intervention. It also provides a comprehensive overview of contemporary arguments with respect to the importance of early childhood intervention on at-risk children and their families throughout the life course. The adolescent life phase is a particular focus. This discussion concludes by highlighting the non health-related outcome domains that are considered significant during adolescence.

Chapter 4 follows directly on from Chapter 3 by discussing, in detail, methods available for measuring quality of life and non health-related quality of life. Further, this chapter outlines the methodology adopted in this thesis to identify non health-related quality of life outcomes domains that are associated with early childhood intervention during adolescence.

#### **1.4.2: Section B (Chapters 5-6)**

Section B develops a method for estimating the effectiveness (in terms of effect size) of early childhood intervention programs on outcomes occurring during adolescence. Effect sizes are calculated from outcome data of longitudinal research, and meta-analysed. A meta-analysis is essentially a statistical method for combining results from a series of studies, whereby the aggregated results provide an estimate of the relationship between variables, or address a series of research hypotheses (Lipsey & Wilson, 2000). Section B is important because it provides the foundation for measuring preferences associated with outcomes during the adolescent life phase with respect to early childhood intervention programs. Preference scores provide rankings of alternatives for program desirability with respect to their perceived contribution to increasing non health-related quality of life. These scores also provide the foundation for measuring the cost-utility of early childhood intervention, ensuring that qualitative improvements in an individual's life can be evaluated from an economic perspective. Moreover, the identification of outcome domains and their indicators is the foundation for developing hierarchies for the purpose of assisting policy-makers in making structured decisions under uncertainty.

Chapter 5 provides an outline of longitudinal research on the effectiveness of early childhood intervention on outcomes during adolescence and introduces the meta-analytic framework to be used in this thesis.

Chapter 6 provides the results of the meta-analysis and discusses the effects of five forms of early childhood intervention (structured preschool program, home visitation, centre-based childcare/developmental day care, family support services and parental education) on at-risk children and their families with respect to seven outcome domains (educational success, cognitive development, social-emotion development, deviancy, social participation, criminal justice outcomes, and family wellbeing) throughout adolescence (age 12-19 years).

### **1.4.3: Section C (Chapters 7-10)**

Section C begins by discussing the theoretical and philosophical foundations of the economic analysis component of the thesis, focusing on alternate methodologies to assess the economic effectiveness of early childhood interventions. This includes a detailed discussion of utility and preference values. A structured methodological procedure is identified for potential use when making complex multi-criteria decisions with respect to policy options for early childhood interventions and capturing preference values for the various alternative forms of early childhood interventions, their outcomes and their indicators during the adolescent life phase. A pilot empirical study was conducted using this new methodology, and the results of this pilot are analysed. Methods of measuring the economic effectiveness of early childhood interventions are examined. On this basis, we examine the concept of

measuring, from an economic perspective, the utility derived from early intervention projects on salient life outcomes (non health-related) over an individual's life course.

Section C is comprised of the following chapters:

Chapter 7 discusses the foundations of economic analysis, focusing on alternate methodologies to assess the economic effectiveness of early childhood interventions. The chapter also includes discussion of possible methods to assist in making more structured decisions with respect to policy decisions regarding early childhood intervention programs and their potential outcomes at any given life phase (e.g. early childhood, adolescence). Moreover, we discuss a method to assist policy-makers in determining the economic costs and benefits of their decisions based on a cost-utility framework. In this chapter, a discussion regarding methods of measuring the cost-utility of crime prevention/early childhood intervention program are included. In particular, we focus on how preference values may be used to measure the economic impact of early childhood intervention programs on qualitative improvements in non-health related quality of life.

Chapter 8 discusses, in detail, the analytic hierarchy process (AHP). First, the issue of organising complex structures into a systematic problem-solving procedure is examined. This incorporates the use of hierarchies in modelling decisions under uncertainty. Issues regarding decomposition (the structuring of a hierarchy to capture all the elements of a given problem), comparative judgments (the development of a matrix to perform pair-wise comparisons of the relative importance of the proposed elements) and the synthesis of priorities (calculating estimates of the vector,

calculation of the consistency index and random index) are discussed. The benefits and criticisms of the AHP method are summarised and an example of its use is provided.

Chapter 9 outlines the method used in this thesis to attain preference values for decisions regarding options for early childhood intervention programs and their contribution to increasing non-health related quality of life during the adolescent life phase. It also offers a methodology for analysing complex multiple criteria problems with the aim of providing a systematic decision making tool to aid in making better policy decisions with respect to choosing program alternatives that have the potential to improve the lives of at-risk children and their families.

Chapter 10 provides the results of a survey that tested the robustness of the AHP method when applied to assist in policy development in the area of early childhood intervention/crime prevention. Moreover, we demonstrate the usefulness of the AHP method in attaining preference values needed for measuring the economic impact of qualitative improvements in non health-related quality of life resulting from early childhood intervention programs during adolescence.

Chapter 11 summarises the findings from the various components of this thesis. Limitations of the research underlying these findings are discussed, and implications for future research are identified. This discussion incorporates a detailed examination of how relative utilities may be utilised in the measurement the cost-utility of early childhood intervention programs. Finally, some concluding remarks regarding the policy implications of this thesis are provided.



Results of Section B (meta-analysis of longitudinal research on the benefits of early childhood intervention programs on adolescent outcomes) (Chapters 5-6) demonstrated that early childhood intervention programs have positive lasting effects on seven outcome domains through late adolescence continuing into early adulthood. Outcomes include (from largest to smallest effect): educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes, and family well-being. The meta-analysis also revealed that early intervention programs that incorporate a structured preschool or centre-based educational component provide lasting results (to the adolescent years and beyond) on the outcome domains educational success and cognitive development. Programs that contained a follow-through component (e.g. preschool continuing up until grade 3) revealed strong effects on the outcomes educational success and cognitive development throughout adolescence. Additionally, the length and intensity of an intervention program were found to be important moderators of success throughout the adolescent life phase. For example, programs whose duration was  $> 3$  years demonstrated significant effects on the outcome domains educational success and cognitive development. Moreover, a program whose intensity was  $> 500$  sessions demonstrated positive results on the outcome domain educational success.

We found that gaps still exist in our knowledge and understanding of developmental prevention and its effects into, and beyond the adolescent years. We posit that ways must be found to customise programs to suit variations in ethnicity to make them more beneficial to the target groups (Homel et al., 2006; The Developmental Crime Prevention Consortium, 1999). Moreover, as argued by McLoyd (1998b), researchers need to evaluate the effectiveness of lengthy and

intensive interventions on children from different ethno-racial backgrounds, and understand the effects of early intervention, beyond the childhood years, on community-level outcomes.

Section C of the thesis (Chapters 7-10) revealed the relative preference and strength of seven outcome domains (educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes and family wellbeing) with respect to their potential contribution to increasing non health-related quality of life during adolescence. Results highlight that the outcome family wellbeing is considered the highest priority during adolescence, followed in order of perceived preference by the outcomes social-emotional development, social participation, criminal justice outcomes, deviancy, educational success, and cognitive development.

Results of Section C also demonstrated that structured preschool programs are considered the highest priority with respect to achieving a high effect on all outcomes during adolescence. This is followed in order of priority by family support services, parental education, centre-based childcare/developmental day care, and home visitation. Additionally, an analysis of results demonstrates that structured preschool programs are considered three times more important than home visitation, and slightly less than three times more important than centre-based childcare/developmental day care with respect to achieving large effects on all outcomes during the adolescent years. Family support services and parental education programs are considered almost equal with respect to their contribution to achieving large effects during adolescence,

which are seen as more than twice as important as the program alternatives centre-based childcare/developmental day care and home visitation.

In addition to providing a method for eliciting relative utility values and a structured method for making decisions regarding policy options for early childhood intervention program alternatives, we have outlined a structured method for use in conducting a cost-utility analysis of early intervention programs (Chapter 11). This incorporates a detailed discussion regarding the concepts of collecting cost data, relative utility data, and the process of combining both costs and utilities to elicit cost-utility ratios. Such results, as discussed earlier in this chapter, provide decision makers with a set of additional tools for making better decisions, which in turn, promise to provide those most at need (high risk populations) with the necessary resources to have a positive effect on the domains of both individual and family functioning throughout the life course.

## **CHAPTER 2: AN ECONOMIC MODEL TO AID IN POLICY DECISION-MAKING AND MEASURING THE COST-UTILITY OF EARLY CHILDHOOD INTERVENTION**

### **2.1: Introduction**

This chapter summarises the crime prevention literature highlighting the need to develop an economic model that provides a structured process for making complex decisions under uncertainty with respect to policy options for early childhood intervention/crime prevention programs. The term early childhood intervention is used to incorporate crime prevention programs. It is recognised that most forms of early childhood intervention aim to direct at-risk children onto positive developmental pathways, which may ultimately lead to improvements in educational success, cognitive development, social-emotional development, social participation, and family wellbeing; as well as improvements in health, reduced rates of delinquency and a reduction in future criminal behaviour. We also outline the need to develop a method that identifies a common metric or set of common metrics that accurately value the qualitative improvements in an individual's life from a cost-utility perspective. Identifying a 'common metric' allows one to investigate the relationship between study features and study outcomes across a variety of programs. By coding the study features according to the objectives of the study, it is possible to transform the outcomes to a common unit so that outcomes of various programs can be compared. It is then possible to apply statistical or economic methods to demonstrate the relationships between study features and outcomes across a variety of program options at different periods in the life course.

Section 2.2 examines the current methods applied by economists in measuring the economic feasibility of early childhood intervention programs. In Section 2.3, methodological gaps are identified in the economic measurement and policy decision-making frameworks of early childhood intervention. This section outlines the need to develop a methodology that can be used to evaluate, from an economic perspective, the qualitative improvements in non health-related quality of life resulting from early childhood intervention across an individual's the life course. A discussion regarding the difficulty of making structured complex multi-criteria decisions under uncertainty, and developing an ordered process to assist in making better policy decisions regarding alternatives of early childhood interventions is outlined. Section 2.4 provides a summary of the chapter and briefly introduces Chapter 3.

## **2.2: Current methods for economic analysis of intervention programs**

Longitudinal research into early childhood intervention has found that well-designed programs provide benefits to the child (e.g. gains in intellectual and achievement scores), the family (e.g. better family functioning) (Brooks-Gunn et al., 2003, p.5-9), and the community (e.g. increases in social capital and collective efficacy) (Commonwealth Taskforce on Child Development Health and Wellbeing, 2007). Additionally, they generate net financial savings to the government (Aos, 2003; Aos et al., 2004; Welsh, 2001). Savings to the government include: reduction in costs to the criminal justice system; reductions in costs borne by victims of crime; and decreases in welfare assistance and a fall in children requiring special education services (Aos, Phipps, Barnoski, & Lieb, 2001; Karoly et al., 1998; Welsh, 2001). For example, a cost-benefit analysis of the Perry Preschool Program yielded a US\$7.16

return per dollar spent (Barnett, 1993). Reduced participation in crime provided the biggest ‘bang for the buck’, with an estimated saving of US\$49,044 per participant. This was followed by increased taxes on earnings (26%), a decrease in the need for special education for participant children (25%), and reduced welfare assistance (9%). An updated cost-benefit analysis of the High/Scope Perry preschool Program, using data on individuals aged 40 revealed the net present values for participants, the general public, and society. Results demonstrate that the treatment group obtained significantly higher earnings compared to those who did not receive the intervention. For the general public, higher tax revenues, lower criminal justice system expenditures, and lower welfare payments easily outweighed program costs. Results reveal that every dollar invested yielded a US\$12.90 return, with program gains coming mainly from reduced crime by males (Belfield, Milagros, Barnett, & Schweinhart, 2006). An economic analysis of the Elmira Prenatal/Early Infancy Project for high-risk families produced similar positive results, with estimated net savings to the government in the order of US\$24,694 per participant (Greenwood, Karoly, Everingham, Hoube, Kilburn, Rydell, Sanders, & Chiesa, 2001).

Although analyses have produced findings demonstrating the benefits of early childhood intervention, there is a debate regarding the optimal timing of such interventions (Brooks-Gunn et al., 2003; Campbell & Ramey, 1994; Currie, 2000). These debates relate to the timing of interventions in order to generate the greatest impact on both health and non-health related quality of life during an individual’s life course, as well as the timing of investment by government so as to gain the most in terms of net financial savings from each dollar invested. This is particularly salient given the competitive nature of funding demands and the need to provide funding

with confidence that the investment will produce positive life changes and opportunities for those at greatest risk. Further, resources available must be able to extend across many sectors of the community; for example health, welfare, education, and crime prevention. Discussion regarding the optimal timing of intervention to reduce negative life course pathways is provided in Chapter 3. For the remainder of this chapter we focus our attention on prominent methodological gaps in the economic evaluation literature and decision-making frameworks relating to choices between early childhood intervention programs.

To date, economic analyses, namely cost-savings, cost-effectiveness and cost-benefit analysis tend to focus on either future savings in governmental or institutional spending or the effectiveness of interventions based on ‘traditional’ cost-effectiveness methods. Cost-savings analysis (CSA) determines the costs and benefits that are realised by a program’s funding body. This type of analysis determines whether the investment ‘pays for itself’. It is common practice that only costs to the funding body are taken into account (Greenwood et al., 2001). Cost-effectiveness analysis (CEA) typically refers to the evaluation of a program and its alternatives with respect to identifying the costs and effects of producing a given outcome (Levin & McEwan, 2001). For example, typical evaluations focus on a choice of a particular intervention compared with its alternatives with respect to a given objective (e.g. increasing academic test scores). Results of the program and its alternatives can then be assessed according to their effects on a given objective (e.g. comparison of test scores at a one year follow-up). With comparable rates of effectiveness between the program and its alternatives and their associated costs, we are able to develop a cost-effectiveness (CE) ratio for comparing the alternatives. Thus, cost divided by the difference in

levels of effectiveness produces a CE ratio. Cost-benefit analysis (CBA) is a systematic and comprehensive cataloguing of costs and benefits that are valued in dollars. Net benefits are then determined relative to the status quo. CBA, when used as a policy assessment tool quantifies in monetary terms the value of policy decisions to all members of society. This is commonly termed 'net social benefits' (NSB), where NSB equals social benefits (B) minus social costs (C) (Boardman, Greenberg, Vining, & Weimer, 1996).

Another popular method, used widely in health and welfare economics is cost-utility analysis (CUA). CUA is essentially a way of describing the relative preference strength of each outcome over a range of health and non-health related domains with respect to their associated costs. In other words, one evaluates program alternatives according to a comparison of their costs and perceived utility (Boardman et al., 2001). The method is similar to that of CEA whereby a ratio is developed (in this case, costs divided by utility). The methods (CEA and CUA) are identical on the cost side but differ significantly on the effectiveness side. As stated above, levels of effectiveness are limited to programme-specific outcomes (e.g. increasing academic test scores) in CEA. However, in CUA, outcomes may be single or multiple, are general as opposed to program specific and may incorporate a notion of value to both the individual and society (Drummond et al., 1997). In CUA, individuals are asked to express their satisfaction with single or multiple measures of effectiveness. Once costs and utilities of the various outcomes are determined, one can choose the program alternative that provides the highest utility at the lowest cost (Levin & McEwan, 2001).



The term ‘utility’, as employed by economists, refers to the satisfaction derived by individuals from one or more outcomes (e.g. improvements in educational success and family wellbeing) (Gold et al., 1996b). The term ‘value’ refers to an individual’s relative preference among the elements, whereby strength of the value is defined by using a set of real numbers (Isard & Smith, 1982). Used widely in the social sciences, particularly economics, the term ‘preference’ is a concept that assumes a real or imagined choice between alternatives and the possibility of rank ordering alternatives, based on happiness, satisfaction, enjoyment, or the utility they provide (Saaty, 2000). Table 2.1 provides a summary of the various approaches to program evaluation using economic analysis, together with their respective strengths and weaknesses.

Each of these methods of economic analysis has a body of economic theory to back them up. However, as with all methods, there is room for improvement in the light of both new research and practical difficulties. The following section highlights the methodological gaps in the economic evaluation of early childhood intervention and the limited ability of policy makers to make fully informed choices based on unstructured methodological procedures that do not incorporate all salient elements of a problem.

**Table 2.1: Approaches to economic analysis of crime prevention programs**

Type of analysis	Measure of cost	Measure of outcomes	Strengths	Weaknesses
Cost-effectiveness	Monetary value of resources	Units of effectiveness (e.g. improvement in academic test scores)	<ul style="list-style-type: none"> <li>• Easy to incorporate standard evaluations of effectiveness</li> <li>• Good for alternatives with small number of objectives</li> </ul>	<ul style="list-style-type: none"> <li>• Hard to interpret when there are multiple measures of effectiveness</li> <li>• Only useful for comparing up to two alternatives; Outcome measures must be comparable.</li> </ul>
Cost-Benefit	Monetary value of resources	Monetary value of benefits	<ul style="list-style-type: none"> <li>• Can judge absolute worth of a project</li> <li>• Can compare CB results across a variety of projects</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to place monetary values on many salient life benefits</li> </ul>
Cost-savings	Monetary value of resources	Monetary savings resulting from impact of intervention	<ul style="list-style-type: none"> <li>• Good for assessing the savings generated to stakeholders</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to place monetary values on many salient life benefits</li> </ul>
Cost-utility	Monetary value of resources	Units of utility	<ul style="list-style-type: none"> <li>• Incorporates individual preferences for units of effectiveness</li> <li>• Incorporates multiple measures of effectiveness into a single measure of utility</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to arrive at consistent and accurate measures of individual preferences</li> <li>• Cannot judge overall worth of a single alternative, only useful for comparing two or more alternatives</li> </ul>

**Adapted from:** Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis* (2nd ed.). London: Sage Publications.

### **2.3: A methodological gap: The costs of crime**

One area of the literature focuses on the costs of crime, incorporating the anticipation of crime, the consequence of crime and the response to crime. Most of this literature relies on estimates by several leading authors in the field namely Anderson (1999), Brand & Price (2000), Brantingham & Easton (1998), Cohen (1988), Cohen, Rust, Steen, & Tidd (2004), Klaus (1994), Mayhew (2003) and Walker (1996). However, their research tends to either “...endorse or oppose some policy position, usually involving incarceration, based on other arguments or evidence” (Nagin, 2001, p.348). Another area of significant growth over the past decade involves the joint consideration of the costs of reducing crime and the costs of operating the criminal justice system. Programs such as the Perry Preschool Program (Barnett, 1993; Schweinhart, 2004), the Elmira Prenatal/Early Infancy Project (Eckenrode et al., 1998; Olds, 2002) and the Seattle Social Development Project (Hawkins et al., 1999) have been rigorously evaluated from this economic perspective. Rather than acting as a ‘rhetorical tool’, as Nagin (2001) views the costs of crime literature, the latter form of economic analysis provides the possibility of making an impact on the formulation of crime control policy.

Little work has been done on developing economic analysis methods in crime prevention beyond the approaches highlighted in Table 2.1. As a result of this dearth of methodological development, Daniel Nagin proposed several strategies that would facilitate growth in the area (Nagin, 2001). The benefit of Nagin’s work is that it focused on the theoretical and practical obstacles that one faces when estimating the benefits of developmental prevention programs. Rather than criticising the more technical matters of economic analysis (e.g. choosing the right discount rate), his

focus is on the broader conceptual issues concerning the overall structure of an analysis. Consequently, what Nagin proposes has the potential to demarcate analyses that are merely rhetorical tools from those that can be policy changing. Nagin's seminal paper (Nagin, 2001) offers insight into where to start looking, since it highlights the need to develop a methodology that measures benefits across multiple domains, at different times, yet at the individual level, rather than for society as a whole.

Traditional economic analyses of crime prevention programs have generally focused on the crime control perspective and concentrated much of their attention on potential savings to government, such as future reductions in costs to the criminal justice system, reductions in costs associated with victims of crime, decreases in welfare assistance, and a fall in children requiring special education services (Aos et al., 2001; Karoly et al., 1998; Welsh, 2001). Recent theorists have argued, however, that this perspective does not suit the developmental prevention perspective, whose natural unit of analysis is the individual (Nagin, 2001). Nagin argues that three changes need to be made when addressing the appropriate unit of analysis: "Individuals or crimes? Society or government? The crime rate and its social consequences or the criminal event and its consequences for the victim?" (Nagin, 2001, p.347). The following sections (Sections 2.3.1-2.3.3) provide an overview of Nagin's main points

### **2.3.1: Individual benefits versus crimes averted?**

Concerning the first change, 'individuals versus crimes', it is argued that a more holistic, individual level approach should be adopted "...one that values benefits

across multiple domains of individual functioning” (Nagin, 2001, p.347). Nagin’s review of the alternative analytical strategies for performing economic analyses of early intervention, namely cost-effectiveness and cost-benefit analysis, highlights the need to adopt a new approach. Rather than focusing solely on the crime control perspective, Nagin argues that the natural unit of analysis in measuring the economic effectiveness of early childhood intervention should be on the individual. That is, the qualitative improvements in a child and his/her family’s quality of life as a result of an intervention should be the focus of the analysis.

Nagin’s main criticism regarding past cost-effectiveness studies, for example these of Greenwood, Model, Rydell & Chiesa (1996) and Donohue & Siegelman (1998) is that the analyses rest too heavily on plausible but highly speculative estimates of the impact of early intervention on crime. The Greenwood et al. (1996) study focused on the crime control effectiveness of California’s Three Strikes statute compared to two forms of early childhood intervention: home visit/day care and parental training programs. The Three Strikes laws are statutes enacted by state governments in the United States, which require the state courts to hand down a mandatory and extended period of incarceration to habitual offenders who have been convicted of a serious criminal offence on three or more separate occasions (Austin, Clark, Hardyman, & Henry, 1999). The analysis of Greenwood and his colleagues suggests that home visit/day care is less cost-effective than California’s Three Strikes statute in preventing serious crime. The cost-effectiveness of the Three Strikes statute was found to be US\$13,899 per serious crime averted compared to the home visit/day care option at US\$89,035. However, the cost-effectiveness of parent-training was only US\$6,351 per serious crime averted. Consequently, Greenwood et al’s analysis was

supportive of parent training but not of home visits/day care. Moreover, the analysis performed by Donohue & Siegelman (1998) proposing that early prevention is a cost-effective alternative compared to imprisonment cannot be convincingly sustained.

This is supported by Nagin who states: "...cost-effectiveness based on a crimes-averted metric is an insufficient evaluation criterion" (Nagin, 2001, p.358). This position derives from the fact that the analysis tends to rest on seemingly plausible, but still highly speculative estimates of the impact of early prevention on later criminality (Nagin, 2001). Similarly, the analysis performed by Greenwood et al. (1996) demonstrating that parent training is a cost-effective alternative to California's 'Three Strikes' law can also be argued to be inherently speculative, given that their findings depend on the extrapolation of the impact of parent-training programs on adult crime. Moreover, "...estimates of the crime control impact of California's Three Strikes law was based solely on crimes averted through incapacitation. If the law also had a deterrent effect, the estimate of the cost-effectiveness of the Three Strikes law was understated" (Nagin, 2001, p.358).

Two important issues are raised by Nagin's review of recent applications of cost-effectiveness methods. First, the current structure of cost-effectiveness analysis does not to date account for the valuable impacts of developmental prevention programs, for example increased public safety, or salient individual benefits, such as improved quality of life. Secondly, there seems to be disjunction between the timing of the initial investment and the future realisation of benefits. Nagin states; "It seems misguided to frame the argument for developmental prevention in narrow crime prevention terms when in fact developmental prevention is not directly competing for

crime control resources”( p.361). Further, Nagin posits that cost-benefit analyses of the Perry Preschool program performed by Barnett (1993), Barnett (1996) and Schweinhart, Barnes & Weikart (1993) failed to capture the full effects of developmental interventions. Why is this? Nagin’s reasons stem from the discussion by Karoly et al. (1998) on the importance of early childhood.

Research and clinical work have found that the experiences of the infant and young child provide the foundation for long-term physical and mental health as well as cognitive development...the period of early childhood development is thus unique-physically, emotionally, and socially. It is a period of both opportunity and vulnerability (pp.2-3).

Although early childhood is a time of opportunity and vulnerability, the importance of opportunities in other life phases and the vulnerabilities these phases may hold should not be discounted. Rather, we should view early childhood as an important phase because it comes first, and not because everything depends on what happens in this early phase. Rutter (2007) states:

...Much credence was placed on the notion of ‘critical periods’ – the assumption being that the early years provided a special ‘window of opportunity’ because of the plasticity of brain development at that age...It is true that brain development undergoes particularly dramatic changes in these early years, but both animal studies and human studies have shown that experiences in adolescence and adult life can make a major impact (Laub and Sampson, 2003). The early years are particularly important, not mainly because of any critical period of effect, but because they come first, and because early experiences tend to shape later experiences (p. 199).

Karoly et al. (1998) conclude that overall, developmental programs as a consequence of their actions provide "...gains in the emotional or cognitive development...improvements in educational process and outcomes...increased economic self sufficiency...reduced levels of criminal activity, and improvements in health-related outcomes" (p.xv). Consequently, it is argued that programs that have potentially far-reaching impacts must be valued in holistic terms rather than in terms of the traditional crime savings perspective (Nagin, 2001, p.364). It is further contended that the broadening of the sampling of outcomes in the calculation of benefits will be ineffective; for example, inclusion of medical costs for treating accidents, and a reduction in costs for child protective services. Nagin (2001) attributes the ineffectiveness of such a method to the expense of collecting such data in a way which would enable demonstration of discrete outcomes. He states, "...many potential impacts such as improved performance in the labour market or lower criminality require years to follow-up to document" (p.364). Moreover, the highly speculative nature of valuing discrete impacts, for example placing monetary values on discrete events such as an arrest or a year of special education services have not been adequately critiqued (Nagin, 2001). Nagin posits that such cost estimates are highly imprecise and that simply increasing the list of items valued only increases the speculative nature of the analysis. Karoly et al. (1998) concurs with Nagin by stating: "...this apparent approach still does not begin to capture the far-reaching effects of an effective intervention" (p.365).



### **2.3.2: Society or government?**

Nagin also questions whether society or government is the correct unit of analysis. Nagin (2001) argues for the first alternative. Although various studies have moved beyond assessing only the costs and benefits of crime reduction (Aos, 2003; Karoly et al., 1998; Olds, Henderson, Phelps, Kitzman, & Hanks, 1993), their perspective has generally remained focused on the benefits of reduced government spending. “Specifically, these studied whether programs had a net positive or negative impact on government expenditures” (Nagin, 2001, p.371). Nagin suggests that this is not necessarily a bad perspective, however using the government perspective as a basis for arguing the case for developmental prevention programs is inappropriate, given its narrow view. Of greater importance is recognition that the government or fiscal perspective undervalues, or is unable to value, the potential benefits to society at large such as increased public safety, or salient individual benefits, such as improved quality of life (Nagin, 2001). Gold, Siegel, Russell & Weinstein (1996b) argue that this problem can be avoided if one takes a broader societal perspective when conducting cost-benefit analyses.

### **2.3.3: The crime rate and its social consequences or the criminal event and consequences to the victim?**

Nagin’s third key point is that when estimating the costs of crime, analysts should value tangible consequences to both victims and non-victims. Traditionally, cost analyses have relied on estimating various components of crime (e.g. physical injury or death, lost or damaged property, and the intangible costs of violation), lost

wages, victim medical fees, and costs to the criminal justice system such as policing and prison expenditures. Certain intangible items have also been estimated such as pain, suffering and loss of quality of life (to victims). Although these are important to measure, the methodology used to value some of these intangibles has been met with skepticism (Nagin, 2001). Cohen (1988) and Cohen, Miller & Rossman (1994), for example, when estimating the costs of individual crimes based mainly on jury awards for pain, suffering and reduced quality of life, did not manage to capture the full effects of crime on society. This was due to the use of *ex post* (after the fact) analysis, rather than the *ex ante* (beforehand) measures of willingness-to-pay preferred by most economists.

Nagin (2001) argues that neglecting non-monetary impacts essentially dismisses the value, or fails to capture the full cost of intangibles that must be incorporated into any more thorough economic analysis. “Listing of the economic costs of crime is incomplete because it neglects non-monetary impacts. Fear of victimization diminishes quality of life by causing individuals to alter their routine activities in otherwise undesirable ways” (Nagin, 2001, p.373). Since Nagin’s article, Cohen et al. (2004) have applied a new methodology to estimate the costs of crime. Their method is based on the contingent valuation method used in environmental economics, which allows estimation of the public’s willingness-to-pay for improvements in non-market goods (e.g. improvements in air quality). Cohen et al.’s latest method does address some of the problems identified by Nagin (2001), by moving beyond assigning costs only to individual victims, to incorporate external social costs to non-victims, neighbourhoods and society in general.

Although Nagin proposes three changes concerning the unit of analysis, we will focus primarily on two, namely, ‘individuals versus crimes averted’ and ‘society or government’. I concur with Nagin that the natural unit of analysis in measuring the economic effectiveness of early childhood intervention is the individual or their family. Consequently, the aim of this Section C of this thesis is to develop a method for measuring salient improvements in quality of life resulting from early-in-life interventions.

#### **2.3.4: A second methodological gap: Making complex funding decisions under uncertainty**

There is not much research in crime prevention literature demonstrating how policy-makers formulate complex multi-criteria decisions under uncertainty. Typically, policy decisions are made in the hope that all relevant criteria are incorporated in the decision-making structure. Moreover, from an economic perspective, there has been little research to identify how funding should be allocated across the various alternatives available (Manning, 2004; Manning et al., 2006). Without a structured procedure for making choices between competing alternatives, which incorporates all relevant criteria, making decisions of this nature is difficult and often inaccurate. We argue that current methods of making funding decisions with respect to early childhood intervention are flawed. Although policy decisions are often made with regard to empirical research, we argue that complex decisions of this nature cannot be made without a structured methodological process, given the complexity of the problem and the large number of criteria that must be incorporated into any decision (e.g. type and makeup of program, effectiveness, number of program components, length and intensity and geographical location etc.). Thus, the

aim of this thesis is to provide a systematic process in which decisions can potentially incorporate all relevant criteria with respect to potential outcomes resulting from early childhood intervention programs. The adoption of this method would allow policy-makers to make better and more informed choices regarding the development of policies to improve the lives of at-risk children and their families (Saaty, 1990).

## **2.4: Conclusion**

The dilemma faced by economists when attempting to place an economic value on early childhood interventions has been the identification of a common metric that accurately values a qualitative improvement in an individual's life. In this chapter, we have examined the disputes surrounds the methods employed in previous cost-benefit analyses. Moreover, we have probed the theoretical and pragmatic obstacles one faces when estimating the economic effectiveness of early childhood interventions. We have argued the need to develop a method that identifies a common metric or set of common metrics that permit valuation of the qualitative improvements in an individual's life from a cost-utility perspective.

Further, we have argued that policy-makers' current decision-making is often flawed. This imperfect model is due to decisions not incorporating complex multi-criteria components into the decision-making framework. To address this gap, we propose a structured method of decision-making, namely an adaptation of the Analytic Hierarchy Process (Saaty, 1990). This economic and decision-making method, we argue, could be applied to evaluate any developmental/crime prevention initiatives. For the purpose of this thesis, however, early childhood intervention programs are used to demonstrate the usefulness of this methodology when our focus

is restricted to outcomes during adolescence. Extension to other life phases is, of course, possible given more time and resources.

In Chapter 3, we discuss, in detail, the literature with respect to crime prevention, early childhood intervention and the developmental perspective. We provide a comprehensive overview of the contemporary arguments with respect to the importance of early childhood intervention on at-risk children and their families throughout the life course. Moreover, we focus in significant detail on the adolescent life phase. Chapter 3 provides the foundation for Chapter 4, which proposes a methodology for identifying non health-related quality of life outcome domains during adolescence and their respective indicators associated with early childhood interventions.

## **CHAPTER 3: PREVENTION, EARLY INTERVENTION AND THE DEVELOPMENTAL PERSPECTIVE**

### **3.1: Introduction**

In Chapter 1 we highlighted the benefits, based on longitudinal research, of well-designed and carefully implemented early childhood intervention programs. These benefits are particularly salient for Australian children and their families who are at greatest risk due to social isolation and economic disadvantage. Moreover, current research demonstrates that there have been few improvements for Australian children over time with respect to health outcomes and child abuse and neglect (Stanley et al., 2005). However, empirical research has produced evidence that intervening with those most at risk early in the developmental pathway can successfully reduce future health, educational, behavioural and crime-related problems (Farrington & Welsh, 2002).

This chapter examines the crime prevention and early childhood intervention literature with a focus on the developmental perspective. Section 3.2 provides a comprehensive synopsis of the contemporary arguments with respect to the importance of early childhood intervention with at-risk children and their families throughout the life course. This includes a discussion on the importance of early childhood intervention and its effect on those most at risk. Section 3.3 focuses attention on the developmental pathways model and introduces the notion of the need to carefully consider the impact of early intervention on various life phase outcomes. Section 3.4 shifts attention to a review of longitudinal studies evaluating the effectiveness of early intervention programs. Section 3.5 analyses the adolescent life

phase. In this section, adolescence is defined together with a discussion regarding transition points, interrelated conceptual domains of risk and protective factors, and adolescent outcome domains. Finally, Section 3.6 provides concluding remarks and briefly introduces Chapter 4.

### **3.2: Prevention and early intervention**

Research results indicate that children who display early signs of disruptive behaviour, who are unprepared for school, or who live in a disadvantaged home environment are at a greater risk of developing social and psychological problems later in life (Golly, Stiller, & Walker, 1998; Homel, Elias, & Hay, 2001). Prevention and early intervention offers at-risk individuals the potential to overcome an array of problems associated with development, health, learning, behaviour and wellbeing over both the short- and long-term (Hayes, 2006). Although one must acknowledge a subtle difference exists between prevention and early intervention. Prevention involves planning and organising efforts to reduce negative pathways while preserving or increasing positive pathways for individuals (Little, 1999). Early intervention on the other hand, a subset of prevention, begins by identifying individuals who are seen as members of a group who are at risk. Early intervention groups may be classified according to whether efforts are placed on the whole population (universal prevention), targeted population subgroups identified as at-risk (selective prevention, or early intervention), or individuals identified with a given problem, or at increased risk of its future development (indicated prevention) (Gordon, 1983).

Prevention and early intervention are not only limited to reducing vulnerability and risk. They also aim to enhance protective factors with the goal of enriching the available pathways for an individual (France & Utting, 2005). Typically, this involves the provision of access to experiences and services that compensate for adverse life circumstances, disadvantage, and vulnerability (Hayes, 2006). For example, parenting programs like the Triple P - Positive Parenting program (Sanders, 2003; Sanders, Markie-Dadds, & Turner, 2003) aim to reduce negative parenting behaviours such as harsh, coercive, and inconsistent parenting and enhance a parent's sense of self-confidence by learning positive parenting skills (Hayes, 2006).

### **3.2.1: The importance of early childhood intervention**

Not all children are born healthy, provided with adequate health care, have access to good nutrition, or live in acceptable housing conditions; not all children are born free of disabilities, or are raised by parents who can comfort, nurture, and provide adequate language, literacy, social problem-solving and behaviour management skills. Therefore, early childhood intervention aims to augment a child's development by providing access to experiences and services that support the child and his/her family to assist in shaping later life experiences (Rutter, 2007). As stated in Chapter 2, the early years are particularly important as this period is a starting point from which individual's can minimise the effect of adverse life circumstances, disadvantage, and vulnerability, and provide opportunities for developing new pathways which promote improved quality of life (Meisels & Shonkoff, 2000).

Quality of life may be defined as subjective wellbeing, the latter reflecting the difference between the expectations and hopes of a person and their present



experiences; in other words, the degree to which a person enjoys the important possibilities of his or her life. Possibilities are a consequence of opportunities and limitations each person has in his/her life and reflect the interaction of personal and environmental factors (Diener, Suh, & Oishi, 1997). A conceptual framework developed by Sen and Nussbaum (1992) emphasises the role of functional capabilities of individuals. When evaluating social states in terms of wellbeing, Sen and Nussbaum emphasise the importance of acknowledging the functional capabilities (e.g. individual freedoms including the ability to live to an old age, engage in economic transactions, and participate in political activities) of individuals instead of utility (e.g. happiness) or access to resources (e.g. income and assets). In this context, poverty would be classified as a capability-deprivation. Sen and Nussbaum posit that emphasis should not be placed on how humans actually function, but their capability to function. Capability-deprivation could come about as a result of ignorance, government oppression, lack of financial resources, or false consciousness (Sen & Nussbaum, 1992). Thus, Sen and Nussbaum emphasise the importance of freedom of choice, individual heterogeneity with respect to individual wellbeing and its measurement.

Early childhood should not be considered a critical period as Bruer (1999) advocates. Rather, early childhood should be considered as the first point in a series of important life phases; each of which contain aspects of vulnerability and opportunity (Laub & Sampson, 2003). Rather than suggesting that once past the age of three there is no possibility of brain development and restructuring, we should consider this point as a critical time for learning as suggested by Hockfield and Lombroso (1998). For example, many cognitive and motor skills are gained quickly during childhood, but

after childhood such skills are not mastered as easily. To demonstrate their theory, Hockfield and Lombroso used acquisition of a second language as an example.

If a child learns a second language early in life, both the native and second language are represented in the same cortical region. In contrast, when a second language is acquired in adulthood, a new language centre that is clearly separated from the native language centre is established in the cortex (Hockfield & Lombroso, 1998, p.2).

Hockfield and Lombroso suggest that although this does not explain why children are able to learn a language more easily than an older individual, it does support findings that early experiences do have an effect on brain development.

Solso (1997) posits that although a child's brain is fully-grown by adolescence, restructuring is not impossible. Rather, it is never too late to learn or change, but the process becomes increasingly complex and difficult. Although the brain reaches its full size in adolescence, it "...continues to be malleable throughout an individual's lifetime, although the regions of the brain where synaptic restructuring occurs vary by age" (Solso cited in Karoly et al. 1998:3). McCain, Mustard, Coffey, Comis, Offord, Desjardins, et al. (1999) present similar findings, but they are more sceptical about the possibility of brain development at an older age. They state: "once the critical periods for brain development are passed, providing the child has not experienced extreme neglect, it is possible to develop the brain's capacity to compensate but it is difficult to achieve its full potential" (McCain et al. 1999:6).

The importance of early childhood development cannot be underestimated, given that it is a period of both opportunity and vulnerability, during which children

develop physically, mentally, emotionally and socially (Homel, 2005). Barnett (1996) identifies a number of factors that mediate whether the period of early childhood development is positive or negative, with both biological and environmental stressors having the potential to compromise a child's healthy development (Karoly et al. 1998). Karoly et al. provided the following examples:

Children with a reduced level of parental stimulation or emotional support may also exhibit socio-emotional problems in childhood that are associated with behaviour problems later in life...Children may face resource constraints - due to low family income or inadequate nutrition or health care - that limit their development during this period, with consequences for outcomes in adolescence or adulthood (p.3).

McCain et al. (1999) suggest that there is clear evidence that good early childhood development programs that involve parents or other primary caregivers can influence how they relate to and care for the children, and can vastly improve a child's outcomes for behaviour, learning and health in later life. They also suggest that not only can the programs benefit the child and his/her family they can also benefit all socio-economic groups in society.

### **3.2.2: The impact of early childhood intervention on at-risk children and their families**

Unfortunately, intervention programs are often implemented at a point where the problem (e.g. behavioural) is considered serious (Golly et al., 1998). An intervention then becomes warranted (according to educational policy), and the at-risk or diagnosed child is diverted to what is commonly termed 'remedial intervention'.

However, by the time the at-risk child is identified as requiring intervention, signs of teacher and peer rejection, alternative placements, poor classroom behaviour and learning difficulties are appearing (Patterson, Reid, & Dishion, 1992).

The good news, however, is that early childhood intervention programs (birth to five years) that employ a risk-focused approach can make positive impacts on outcomes of children who are considered at-risk at an early age (Homel, 2005). An influential series of experiments have produced solid evidence that intervening (with the child and the family) early in the developmental pathway is successfully heading-off future health, behavioural and crime problems (Farrington & Welsh, 2002). Evidence is available demonstrating the short- and long-term effects of early childhood intervention projects. 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 et al., 1999) have confirmed the positive effects that early childhood intervention projects have on at-risk children. This is particularly evident in children who come from low- income backgrounds, with "...initial gains in intellectual and achievement scores, and longer-term outcomes reflecting more successful school experiences...reduction of behavioural problems and delinquency have also been reported" (Brooks-Gunn et al., 2003, p.5-9). However, a significant proportion of studies demonstrating short-, but particularly long-term outcomes have been conducted outside Australia, with most coming from the U.S.A. Furthermore, a large proportion of 'disadvantaged' or 'low-income' populations cited in these studies are African American.

Successful experiments have demonstrated that a systematic delivery of basic services or resources to disadvantaged families has resulted in large reductions in crime involvement amongst targeted groups (Reynolds et al., 2004; Schweinhart, 2004; Schweinhart et al., 1993; Yoshikawa, 1994; Zigler et al., 1992). Studies have also confirmed improved outcomes for those targeted across multiple domains including improved educational outcomes, decreases in child maltreatment, reductions in child and youth antisocial behaviour, lower levels of substance abuse, and increases in income and workforce participation (Brooks-Gunn et al., 2003; Olds, 2002; Reynolds et al., 2001). Heckman, Stixrud and Urzua (2006) also stressed the importance of the family in mediating the cognitive and social-emotional skills of children in their early years. Heckman et al's., research highlights that families who do invest in developing their child's cognitive and social-emotional skills (non-cognitive) significantly impact on their child's life trajectory; with reductions in subsequent involvement in crime, teenage pregnancy and educational outcomes. James Heckman's research has highlighted the importance of cognitive and non-cognitive skill development, and the role families play nurturing this development, to overcome social disadvantage in a sustainable manner.

### **3.3: The Developmental Pathways model**

#### **3.3.1: Defining developmental prevention**

The Developmental Crime Prevention Consortium (1999), in contrast to Farrington's (1996) view that developmental prevention is about inhibiting the development of criminal potential in individuals, took a life-course approach that did not assume the existence of propensity to offend (Hemel, 2005). Although the effects

of an individual's life history on the predisposition to offend were not totally discounted, Homel, Elias and Hay (2001) define developmental prevention by stating:

Developmental prevention involves intervention early in the developmental pathways that lead to crime and related problems, emphasising investment in 'child friendly' institutions, communities and social policies and the manipulation of multiple risk and protective factors at different levels of the social ecology and at crucial transition points, such as around birth, the commencement of school, or graduation from primary to high school (p.272).

The developmental pathways focused on in this thesis are distinguished from the more causal pathways models proposed by Hertzman (1999). These are written from a medical or epidemiological perspective where there is an emphasis on "linear causal chains of events" (Homel, 2005, p.83). For example, one condition (e.g. poor school readiness) leads to another (e.g. disability and absenteeism in the fifth decade of life) via intermediate events (e.g. working in a highly stressed, low control job)(Homel et al., 2006). Homel (2005) argues: "According to this model, much of the social gradient effect in health outcomes arises from the amplification and reproduction by social processes of the effects of differences in individual traits and in life circumstances at (or before) birth" (p.83).

The developmental pathways model and the causal pathways model are similar in many points. They do however differ regarding one issue; the developmental pathways model acknowledges that human agency and the possibility of changing one's ways is achievable because of changes in social circumstances and opportunities that arise during an individual's life-course (Elder, 1998; Homel, 2005).

The definition of developmental prevention offered by Homel, Elias and Hay (2001) encompasses all the components that were posited in the 'pathways' report (The Developmental Crime Prevention Consortium, 1999), however, several additional principles must be acknowledged. First, 'early intervention' does not necessarily mean early in life, but 'early in the pathway'. Secondly, it is acknowledged that the context is always changing. This includes, but is not limited to, social policies, institutions and neighbourhoods. Thirdly, it is acknowledged that risk and protective factors do matter. Finally, interventions are highly effective when the focus is placed on transition points. Homel, Elias and Hay (2001) argue:

No one program can cover the waterfront, especially in its early stages, so organising one's thinking around one or at most two key life transitions simplifies the planning task while increasing the chances that interventions will have a high uptake by the target population (p.272 ).

### **3.3.2: Life phases, points of change and transition points**

To this end, Homel (2005) argues that the developmental pathways perspective does not take the view that an individual's life-course is fixed after early childhood. Rather, an individual's life course is marked by a series of phases, points of change, and transitions. For example, the journey to becoming an adult moves through various points of transition; these include, and are not limited to, the transition from home to school, from primary to secondary school, from high school into the workforce, and the possibility of leaving home. Further, transitions during adulthood also occur. These may include shifts in employment status, making

commitments to other people, parenthood, and watching the next generation move through all the phases they experienced (Hemel, 2005).

These changes are sometimes met with apprehension, given the uncertainty of how to make the change. More specifically, most transitions require an individual to identify with new social institutions, many of these requiring an individual to cope with a new set of developmental tasks and challenges (Laub & Sampson, 2005; The Developmental Crime Prevention Consortium, 1999). Moreover, Hemel (2005) argues that points of transition are when interventions are the most effective due to an individual's vulnerability to making bad choices, or "taking false steps" (Hemel, 2005, p.81), and their openness to "external support or advice" (p.81).

Laub and Sampson (2003) posit that the actions of a person in a particular time are dependent on their current circumstances as well as the choices they make and chance events. They also acknowledge the influence of earlier events on behaviour. Hemel (2005) concurs and in particular argues that an individual's current decision framework is dependent to some extent on how they have coped with past problems, supplemented with the extent to which the individual has been equipped with the skills to manage the current problem. Hemel states:

A first time offender, for example, needs to be ready to listen, to feel shame, empathy, and embarrassment. His or her family also need to have developed sufficient will and trust to be able to cope with this particular false step and to move on effectively. Whether they can do so depends on what has happened at earlier points in life. If those earlier situations have led to distrust, alienation, or entrenchment and unproductive strategies for dealing



with difficulty, then success in working through this new problem will be all the more difficult to achieve (p. 84).

### **3.4: Reviews of the literature on the effectiveness of preschool prevention programs**

An extensive literature has developed around life phases and transition points. For example, Ramey and Ramey (1998) have researched the transition from home to school, with a particular focus on 'readiness for school' and how this concept can be understood and improved to aid children in this vulnerable period (France & Homel, 2006; Homel, 2005). In addition a number of narrative reviews highlighting the effectiveness of preschool prevention/early childhood intervention programs exist. Examples include, Cohen and Radford (1999), Karoly et al., (1998), MacMillan, MacMillan, Offord, Griffith, and MacMillan (1994), McCain and Mustard (2002), Nelson, Laurendeau, Chamberland, and Peirson (2001), Ramey and Ramey (1998), Reynolds (1994), Reynolds et al., (2001), Yoshikawa (1994), and Zigler et al., (1992). These reviews have provided insight into the effects of preschool prevention/early childhood intervention programs on various outcome domains over a range of time periods. One conclusion is that early childhood intervention programs that include a direct teaching component (e.g. centre-based preschool education) impact immediately on the cognitive development of children, but gains typically fade when children reach early primary school (Consortium for Longitudinal Studies, 1983; Nelson et al., 2001). The literature also suggests that early childhood intervention programs that focus on children indirectly through parent or family support (e.g. nurse home visitation) are less likely to impact positively on cognitive outcomes than programs with a direct teaching component (e.g. preschool-based educational

activities) (Cohen & Radford, 1999; Ramey & Ramey, 1998). Moreover, programs that provide supplementary academic assistance (e.g. follow through programs) when entering primary school, result in children being more likely to maintain the gains that stem from preschool intervention (Ramey & Ramey, 1998).

Early childhood intervention programs can have positive impacts on many domains of individual and family functioning. Outcomes include reduced rates of delinquency and criminal behaviour, program participants earning higher adult wages, home ownership and fewer arrests than the control group (Berrueta-Clement, Schweinhart, Barnett, Epstein, & Weikart, 1984; Schweinhart et al., 1993; Schweinhart, Montie, Xiang, Barnett, Belfield, & Nores, 2005). Children's social-emotional behaviour and parent-family wellness are other domains which have displayed positive results. For example, the Prenatal/Early Infancy Project follow-up (Olds & Korfmacher, 1998) found that their nurse home visitation program displayed short-term positive impacts on reduced rates of child maltreatment. Over time, those impacts have become more pronounced, particularly for a sub-set of high-risk women. A 15-year follow-up study (Olds, Eckenrode, Henderson, Kitzman, Powers, Cole, Sidora, Morris, Pettitt, & Luckey, 1997) found that mothers who were visited by nurses had higher rates of employment, lower rates of impairments due to alcohol abuse, lower rates of child abuse and neglect, and fewer arrests compared to control group women. Children of the experiment group, at age 15, had significantly lower rates of arrests, were less inclined to run away from home, and had significantly reduced rates of criminal conviction and violations of probation (Eckenrode et al., 1998).

Another important characteristic that moderates effectiveness is program comprehensiveness. Longitudinal research suggests that programs that are multi-dimensional (e.g. programs that focus on children, parents and the community) achieve better outcomes than more narrowly focused programs (e.g. early Head Start programs, parent training or home visitation alone) (Andrews, Blumenthal, Johnson, Kahn, Ferguson, Lasater, Malone, & Wallace, 1982). Literature reviews have suggested that multi-dimensional programs that provide both preschool and home-based support result in positive outcomes on children's cognitive, social-emotional behaviour and family well-being (Yoshikawa, 1994; Zigler et al., 1992). The Parent Child Development Centres (Andrews et al., 1982) and the Pathways to Prevention Project (Homel et al., 2006) are two such programs that have a multi-dimensional focus. These programs consist of comprehensive parenting programs that incorporate child development, child-rearing and home management practices (e.g. Triple P (Sanders, 1999)), nutrition and health in the home, parents' personal development, culturally sensitive playgroups, improved access to government and community resources, and extensive support services for participating families.

The timing, length and intensity of an intervention is another important moderator of program effectiveness. The meta-analytic review of the effect of preschool intervention programs on children's cognitive, social-emotional and parent-family outcomes conducted by Nelson et al. (2003) found that programs less than one year in duration, and fewer than 300 sessions, had minimal impact on children. Previous reviewers of the preschool program literature (Cohen & Radford, 1999; Homel, 2005; Nelson et al., 2001; Yoshikawa, 1994) have also argued that programs

that begin early in a child's life-course may be more effective than programs that begin in the later preschool years.

Another moderating variable consists of the characteristics of children and families enrolled in preschool intervention programs. Lee, Schnurr and Brooks-Gunn (1988) found that ethno-racial background and socio-economic status were moderators of Head Start success. A review of this study by McLoyd (1998a) found that the short-term impacts of Head Start were greater for African-American children than for white children. It was posited that more benefit was gained because of severe economic disadvantage and related environmental stressors (e.g. high rates of poverty, crime and violence in those neighbourhoods), and because African-American children were more likely to live in more dire and prolonged states of poverty (Nelson et al., 2003). Lee et al., (1988) argues:

In general, children who started out the lowest gained the most. African-American children were more likely to be big gainers than White children because they were relatively more disadvantaged demographically and scored significantly lower on all four measures in the prevention year (p.163).

### **3.5: The transition from childhood to adulthood**

As highlighted in Chapter 1, this thesis concentrates its attention on the economic effectiveness of various forms of early childhood intervention (e.g. structured preschool programs (SPP), centre-based child care/developmental day care programs (CBCC), home visitation (HV), family support services (FSS) and parental education (PE)) with respect to their contribution to non health-related improvements in quality of life (e.g. educational success, cognitive development, social-emotional

development, deviancy, social participation, criminal justice outcomes and family wellbeing) during the adolescent life phase. Consequently, it is important to review available evidence relating to the adolescent life phase and early intervention programs.

### **3.5.1: Defining adolescence**

Ages of adolescence vary by culture, but the World Health Organisation defines adolescence as the period of life between the ages of 10 and 19 years (World Health Organisation, 1986). This is in contrast to the United States where the adolescent years typically range between 12 to 20 years of age (Jessor, 1992). More generally, the interpretation of who is considered an adolescent is tied in with the word 'teenager'. This describes a person who is aged thirteen to nineteen years of age (Jessor, 1993; Takanishi, 1993). Others (Eccles, Midgley, Wigfield, Buchanan, Reuman, Flanagan, & Mac Iver, 1993) suggest that it may be worthwhile to consider the starting age as a period of life when most children go through the physical changes of puberty; roughly from 9 to 13 years of age; with the period of adolescence concluding when the transition from high school to the workforce or postsecondary education occurs (approximately 18-19 years of age). Ideally, we should attempt to reduce our reliance on finding adolescent end points that are tied to physical milestones, and consider adolescence as a cultural and social phenomenon that is identified with dramatic changes in one's body, together with developments in individual psychology and academic/work career (Eccles et al., 1993).

In this thesis, we define adolescence by considering the cultural and social phenomenon identified by developments in individual biology and academic standing.

Thus, we use the age range of 12-19 years, which includes the transition to high school and the transition from high school to the work force. This is also a practical choice given a significant number of studies conduct follow-ups in and around these transition periods.

### **3.5.2: Transition points during the adolescent life phase**

The past 25 to 30 years has seen a dramatic increase in the amount of attention being paid to the adolescent years. It is a life phase that is characterised by changes in pubertal development, social role redefinitions, school transitions, cognitive development, and overt signs of sexual maturity (Eccles et al., 1993). Developmental scientists have been drawn to the nature and pace of these changes and the impact that past participation in early intervention programs have on the adolescent years (Lally, Mangione, & Honig, 1988). Acknowledging that adolescence is characterised by significant changes over a large time period (e.g. 8-10 years) that require an individual to identify with new social institutions and develop new coping strategies to deal with new developmental tasks and challenges, researchers have increasingly focused their attention on transitions within this life phase.

Adolescence is marked by two crucial transition points; the transition from primary to high school and the transition from secondary school to either the workforce or further post-secondary education (e.g. academic or vocational training) (Takanishi, 1993). Individuals who pass through these transitions encounter different challenges that impact, either positively or negatively, on how they cope with the stress of these changes (Jessor, 1992). For example, the transition from primary school to secondary school involves adapting to high school, new peer relationships

and different disciplinary structures. The transition from high school to work reveals a new array of stressors, including adult roles and responsibilities, employer expectations, moving away from home, and new economic responsibilities (The Developmental Crime Prevention Consortium, 1999). Table 3.1 provides an overview of the two adolescent transitions, the associated risk factors and possible preventative strategies.

**Table 3.1: Adolescent developmental transitions: Stressors, risk factors and preventative strategies**

Transition	Stressors	Risk Factors	Preventative strategies
Transition to high school	<ul style="list-style-type: none"> <li>-Adapting to high school</li> <li>-Defining identity</li> <li>-Growth of autonomy in a context of peer conformity</li> <li>-Developing value system</li> <li>-Intimate relationships</li> <li>-Different disciplinary structures</li> </ul>	<ul style="list-style-type: none"> <li>-Teenage pregnancy</li> <li>-Risk-taking</li> <li>-Inadequate behaviour management</li> <li>-Unemployment</li> <li>-Antisocial peers</li> <li>-Peer rejection</li> <li>-Lack of parental support</li> <li>-Unprepared academically for high school</li> <li>-Poor attachment to school</li> <li>-School failure</li> </ul>	<ul style="list-style-type: none"> <li>- Anti-bullying programs</li> <li>-Community support for youth in schools</li> <li>-Social skills strategies</li> <li>-Employment guidance</li> <li>-Education support programs</li> </ul>
Transition to workforce or post-secondary education	<ul style="list-style-type: none"> <li>-Adult roles and responsibilities</li> <li>-Employer expectations</li> <li>-Moving away from home</li> <li>-Economic responsibilities</li> <li>-Post-secondary academic expectations</li> </ul>	<ul style="list-style-type: none"> <li>-Unemployment</li> <li>-Poverty</li> <li>-Homelessness</li> <li>-Social isolation</li> <li>-Lack of support services</li> <li>-Norms concerning violence as acceptable response to frustration</li> </ul>	<ul style="list-style-type: none"> <li>-Social and economic development</li> <li>-Building social networks</li> <li>-Work experience programs</li> </ul>

Adapted from: The Developmental Crime Prevention Consortium. (1999). *Pathways to prevention: Developmental and early intervention approaches to crime in Australia*. Brisbane: National Crime Prevention: Commonwealth Attorney-General's Department.

### **3.5.3: Interrelated conceptual domains of risk and protective factors during adolescence**

The characteristic changes that occur throughout adolescence, together with the rapid pace of these changes, has seen adolescence as a time of vulnerability, particularly for at-risk individuals. Consequently, the coping strategies that an individual already possesses, or subsequently attains, during the early adolescent years may impact positively or negatively on an individual's ability to negotiate future challenges (e.g. transition from high school to the work force) (Eccles et al., 1993; Jessor, 1993). Brooks-Gunn, Fuligni and Berlin (2003) observe that most individuals pass through this phase unscathed, but those from socially isolated ethnic and minority groups or low socio-economic backgrounds tend to have the most problems.

Developmental science has evolved to a stage where the development of inquiry continues to expand, allowing greater complexity to be recognised with respect to possible sources of explanation for adolescent risk behaviour (Jessor, 1992, 1993). Jessor (1993) argues that the accumulation of knowledge about adolescent risk behaviour has:

...revealed its intractability to simple explanation-whether focused on a single variable, such as self-esteem, a single setting, such as the inner city neighbourhood, or a single explanatory domain, such as personality, the environment, or genetic disposition...Research in this field has evolved from early descriptive accounts and epidemiological surveys to more and more complex explanations implicating multiple interacting domains that now range from biology to the social environment (p.119).

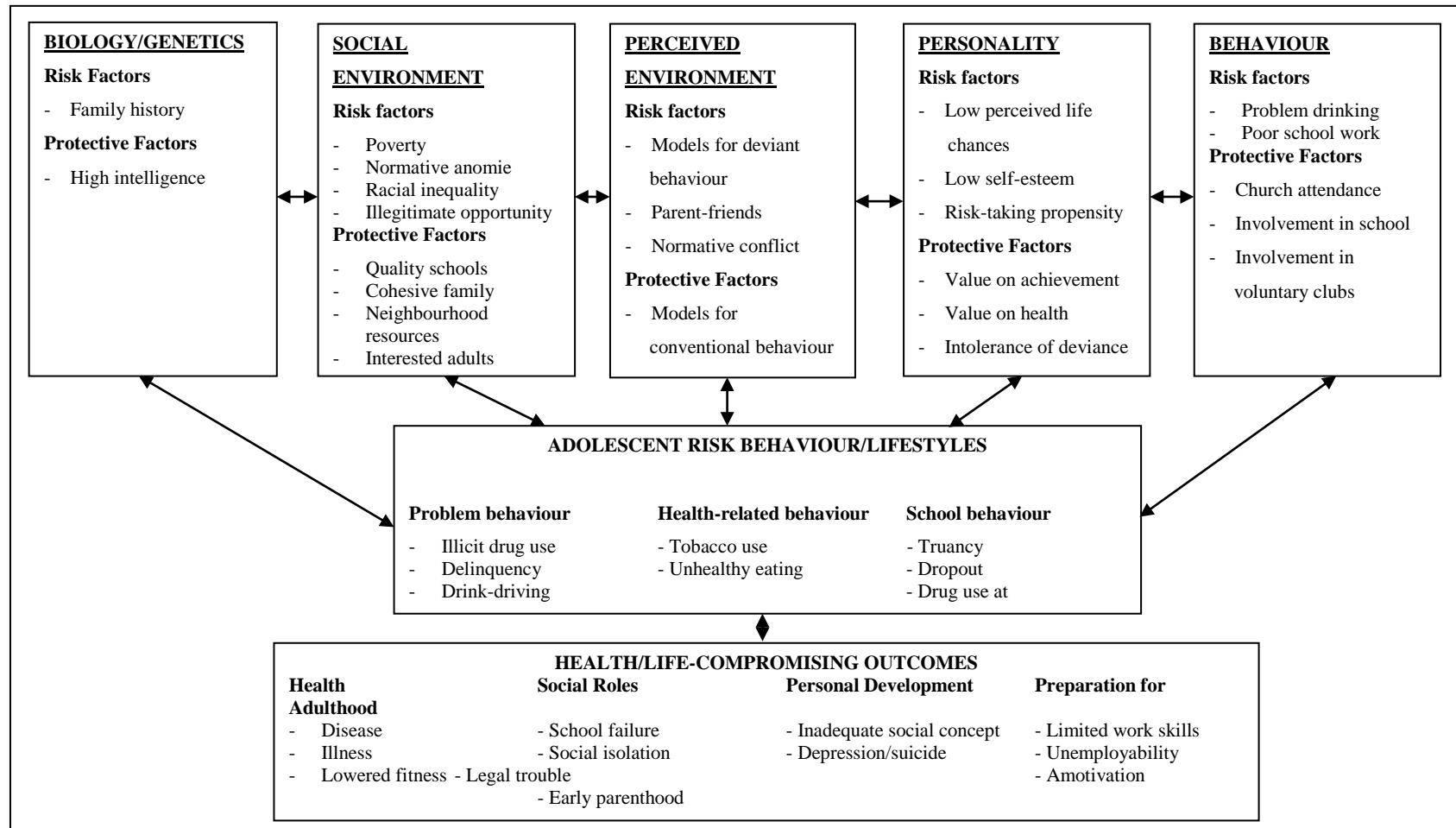


Research into this life phase has seen: increased complexity in multiple explanatory domains; reciprocal and bidirectional causality between domains; differentiation of constructs within domains and their segregation into categories of risk and protective factors; more emphasis on perceived and objective factors; the direct and indirect paths that link the domains with risk behaviours and risk outcomes expanded; risk behaviours expanded into broader life-styles; and, risk behaviours linked to long-term life outcomes (Jessor, 1992, 1993). Jessor (1993) argues: “Understanding contextual change becomes as important as understanding individual change” (p.120). Figure 3.1 illustrates this point by demonstrating that any research attempting to capture variance in adolescent risk behaviour and to understand the role of risk behaviour in development, requires a model that encompasses the interrelated conceptual domains of risk and protective factors (Jessor, 1993).

#### **3.5.4: Adolescent outcome domains**

Longitudinal researchers of early childhood intervention programs, for example, the Elmira Prenatal/ Early Infancy Project (Eckenrode, Ganzel, Henderson, Smith, Olds, Powers, Cole, Kitzman, & Sidora, 2000b; Olds, Henderson, Cole, Eckenrode, Kitzman, Luckey, Pettitt, Sidora, Morris, & Powers, 1998), the Perry Preschool Program (Berrueta-Clement et al., 1984), and the Syracuse Family Research Development Program (FDRP) (Lally et al., 1988) have focused on the main transitions points within the adolescent life phase, highlighting possible outcomes.

**Figure 3.1: A conceptual framework for adolescent risk behaviour: risk and protective factors, risk behaviours, and risk outcomes**



**Source:** Jessor, R. (1992). Risk behaviour in adolescence: A psychosocial framework for understanding and action. In D. E. Rogers & E. Ginzburg (Eds.), *Adolescents at risk: Medical and social perspectives*. Boulder, CO: Westview Press. (p.2)

The outcome domains and their various indicators have typically been derived from domains adopted in previous longitudinal studies, with few longitudinal studies reporting a systematic approach to identifying relevant outcome domains and indicators for their particular population. Identifying outcome domains and their respective indicators for most life phases may be achieved by analysing psychometric test instruments or conducting a detailed review of longitudinal studies on salient outcome domains and their indicators using a meta-analytic approach. In this thesis, we have adopted both methods to identify the relevant outcome domains during adolescence (and other life phases – Appendix A) and their respective indicators. For example, outcomes associated with the transition from primary to high school include cognitive development, social-emotional development, deviancy and social participation, and family wellbeing. The transition from high school to the work force or post-secondary education includes outcomes such as educational success, social participation, criminal justice outcomes, socioeconomic success, social-emotional development, and health. Table 3.2 provides an overview of the various outcome domains and items associated with transitions in the adolescent life phase. In Chapter 4 (Section 4.5) we examine the quality of life literature and highlight the methodology adopted in this thesis to identify the salient non health-related quality of life outcomes domains and their indicators (as seen in Table 3.2) associated with the impact of early childhood intervention on adolescence.

**Table 3.2: Adolescent outcome domains and their indicators**

<b>Educational success</b>	<b>Cognitive development</b>	<b>Social-emotional development</b>	<b>Deviancy</b>
1. Special education 2. Graduation 3. Drop-out 4. Grade retention 5. Completed years of education 6. Absenteeism 7. Reduced learning problems 8. Feeling of belonging at school	1. IQ 2. Achievement tests 3. School grades 4. Rating of academic skill and performance 5. School failure	1. Self-esteem 2. Self-confidence 3. Social skills 4. Parent- teacher rating of problem behaviour 5. Obsessive-compulsive behaviour	1. Delinquent behaviour 2. Running away from home 3. Drug use 4. Lying about age 5. Gang involvement
<b>Social Participation</b>	<b>Criminal Justice Outcomes</b>	<b>Family Wellbeing</b>	
1. Employment in teen years 2. Socioeconomic success 3. Social integration 4. Engaged in skilled labour	1. Rates of juvenile arrest 2. Incarceration 3. Rates of violent arrest 4. Rates of non-violent arrest 5. Petition requests to juvenile courts	1. Child maltreatment 2. Parental-adolescent relationship 3. Family functioning 4. Parental involvement in child's schooling 5. Parental conflict	

Outcomes and relevant indicators of early childhood intervention programs for the adolescent years (ages 13-19 years) were derived from findings of the meta-analysis (Chapters 5 and 6). Early intervention programs that were evaluated during the adolescent years are presented in Table 3.3.

**Table 3.3: Projects and studies included in the meta-analysis, from which outcomes domains and their indicators were developed - Adolescent years**

<b>Project</b>	<b>Study</b>	<b>Outcomes measured</b>
Abecedarian Project	Campbell, Ramey, Pungello, Sparling, & Miller-Johnson (2002)	Educational success, cognitive development, deviancy, social participation, criminal justice outcomes
Abecedarian Project	Campbell & Ramey(1994)	Educational success, cognitive development
Abecedarian Project	(Campbell & Ramey, 1995a)	Educational success cognitive development
Chicago Child-Parent Centre	Reynolds, Temple, Robertson, & Mann,(2001)	Educational success, criminal justice outcomes
Chicago Child-Parent Centre	Reynolds (1994)	Educational success, cognitive development, social-emotional development, family wellbeing
Early Training Project	Lazar, Darlington, Murray, Royce, Snipper, & Ramey (1982)	Educational success, cognitive development, social-emotional development
Early Training Project	Gray & Klaus (1970)	Cognitive development
Elmira Prenatal/ Early Infancy Project	Eckenrode, Ganzel, Henderson, Smith, Olds, Powers, Cole, Kitzman, & Sidora.(2000a)	Family wellbeing
Learning to Learn Project	Sprigle & Shaefer (1985)	Educational success, cognitive development
Louisville Experiment	Miller, & Bizzell (1983)	Cognitive development
Mother-Child Home Program	Levenstein, Levenstein, Shiminski & Stolzberg (1998)	Educational success
Perry Preschool Program	Berrueta-Clement, Schweinhart, Barnett, Epstein & Weikart, (1984)	Deviancy, criminal justice outcomes
The Syracuse Family Research Development Program	Lally, Mangione & Honig (1988)	Educational success, social-emotional development, social participation, criminal justice outcomes, and family wellbeing
Parent-Child Development Centres	Johnson (2006a)	Social-emotional development
Parent-Child Development Centres	Johnson & Blumenthal (2004a)	Educational success, cognitive development, social-emotional development, social participation, and family wellbeing
Direct Instruction project	Meyer (1984)	Educational success, cognitive development
Elmira Prenatal/ Early Infancy Project	Olds, Henderson, Cole, Eckenrode, Kitzman, Luckey, Pettitt, Sidora, Morris & Powers (1998)	Educational success, deviancy, and criminal justice outcomes

### 3.6: Conclusion

In this chapter we have examined the crime prevention and early childhood intervention literature, highlighting that developmental prevention/early intervention offer at-risk individuals the potential to overcome an array of problems associated with development, health, learning, behaviour and wellbeing over both the short- and long-term. In Chapter five, we outline the application of a technique (meta-analysis: a summary of longitudinal studies) for summarising the literature, in particular, the outcomes associated with early childhood intervention and the enduring effects of these interventions on at-risk individuals during the adolescent life phase. Moreover, we outline how the meta-analytic technique also serves as the foundation for the development of a method to measure improvements in quality of life (through the application of a cost-utility analysis) and as a method to assist in making complex policy decisions under uncertainty (Section C of this thesis). Before conducting such an analysis, we need to identify the salient outcome domains during this phase. Constructing outcome domains and identifying their respective indicators allows for a more structured and thorough analysis of the effectiveness of early childhood interventions on a series of non health-related outcomes during the adolescent life phase. Consequently, in Chapter 4, we provide an overview of the quality of life literature, and discuss the methodology adopted in this thesis to identify salient non health-related quality of life outcomes domains and their indicators associated with early childhood intervention during adolescence.

## **CHAPTER 4: QUALITY OF LIFE AND THE MEASUREMENT OF OUTCOMES ASSOCIATED WITH EARLY CHILDHOOD INTERVENTION DURING ADOLESCENCE**

### **4.1: Introduction**

This chapter discusses, in detail, the concept of quality of life. In so doing it introduces the methodology adopted in this thesis to identify non health-related quality of life outcomes domains and their respective indicators associated with early childhood intervention during adolescence. Section 4.2 examines the concept of quality of life (QOL) focusing on both health (HRQOL) and non-health related quality of life (NHRQOL). Further, the domains of QOL and NHRQOL are considered. Section 4.3 discusses the measurement of NHRQOL domains and the mapping of their interactions. Section 4.4 highlights instruments that can be used to operationalise health and non-health related quality of life domains. Section 4.5 discusses an instrument for use to gauge improvements in NHRQOL during adolescence resulting from early childhood intervention programs. This section also examines the concepts of specifying measurement goals, item generation, item reduction and refinement, questionnaire formatting, instrument testing, reliability measurements, testing consistency and conducting sensitivity analysis. Finally, Section 4.6 provides a summary of this chapter and briefly introduces Chapter 5.

### **4.2: Quality of Life (QOL): Health and non-health related**

Measuring QOL has become a salient measure of a program's efficacy and is widely becoming accepted as a valid indicator of a program's merit, particularly in the

health and welfare sector (Spilker, 1996b). Growth in interest can be attributed to qualities such as its easy adaptation to individual, group, or large populations of patients/participants (Gold et al., 1996b). Overall, QOL may be defined as subjective wellbeing, where the subjectivity of QOL reflects the difference between the expectations and hopes of a person and their present experiences; in other words, the degree to which a person enjoys the important possibilities of his or her life. Possibilities are a consequence of opportunities and limitations each person has in his/her life and reflect the interaction of personal and environmental factors (Diener et al., 1997). Humans constantly adapt their expectations such that their expectations lie within the realm of possibility as perceived by the individual. Consequently, this enables individuals who have difficult lives as a result of factors such as low socio-economic status, social isolation, mental and physical disabilities, and cultural isolation, to still maintain reasonable quality of life (Cummins, 2000a, 2000b; Cummins & Lau, 2004).

QOL can be disaggregated into two forms; health-related QOL (being the most common form applied) and non-health related quality of life (NHRQOL) (e.g. social, economic, and religious). The World Health Organisation posit "...health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (World Health Organisation, 1996, p.1). Therefore, a definition of QOL is not necessarily limited to only health outcomes (e.g. physical functioning, absence of disease), and thus evaluations of any form of intervention (e.g. early childhood, secondary intervention) should be broadened to accommodate both health and non health-related outcomes. Nagin (2001) and Gold et al. (1996b) agree that non health-related outcomes, which incorporate life events that are independent of one's health,



should be common practice in the economic evaluation of crime prevention/early childhood intervention programs. It should be acknowledged that total health is a complex tapestry of factors of wellbeing, for example, physical health and safety are a primary dimension grounding a child's overall wellbeing (World Health Organisation, 1996). Additionally, there is a health dimension to wellbeing but in this thesis we clearly distinguish between health states (e.g. physical functioning, absence of disease) and non-health states (e.g. cognitive development, social-emotional development and family wellbeing). A detailed analysis of the psychometric properties of the outcome measures relating to individual non health-related quality of life was conducted. For example, we defined non health-states relating to family wellbeing as strength, unity and improved functioning within the family (Children's Home Society and Family Services, 2007).

Instruments and measures to ascertain NHRQOL improvements resulting from crime prevention/early childhood developmental intervention programs are limited (Gold et al., 1996b; Nagin, 2001). However, researchers and policy makers are now seeing the potential these measures offer. This is not a new phenomenon; health economists have seen the importance of both forms of quality of life, and consequently place health and non-health outcomes on an equal footing. However, the difficulty facing most economists in the area of crime prevention/early intervention is how to accurately measure NHRQOL with the current tools, ensuring that a common set of metric outcomes are established. This difficulty is compounded by variation between policy sectors (e.g. health, welfare, and criminal justice), the various offshoots of these sectors, levels of aggregation of QOL (e.g. the home, community, region, group of nations, and the world), and patient levels of QOL (Garber,

Weinstein, Torrance, & Kamlet, 1996; Gold et al., 1996b; Mandelblatt, Fryback, Weinstein, Russell, Gold, & Hadorn, 1996; Spilker, 1996b). Spilker (1996b) identify five levels of patient quality of life; these include:

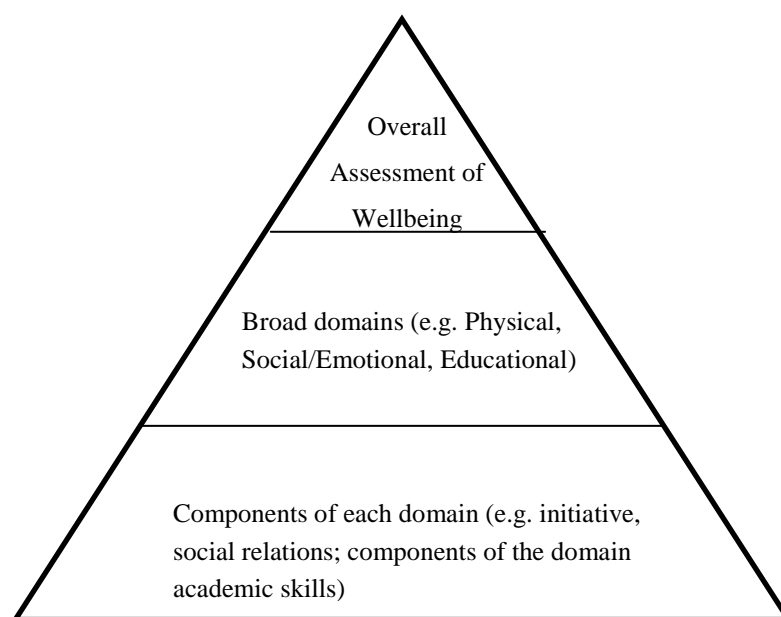
- Level 1: The individual patient/participant;
- Level 2: Patients or participants in a trial;
- Level 3: Patients or participants who have a specific disease subtype;
- Level 4: Patients/participants with selected characteristics who a specific disease or risk factor or a group of risk factors; and
- Level 5: All participants/patients with a specific disease or risk factor.

For the purposes of this study, QOL will be defined as a broad set of life domains that represent the effect of a series of states which impact, either positively or negatively, on an individual's life trajectory. Specifically, this research defines QOL as a set of non health-related domains (e.g. cognitive, social-emotional) specifically related to the outcomes associated with early childhood intervention on adolescence. The domains represent the effect of a series of states (e.g. school achievement tests, social skills, self-esteem, behaviour), which impact, either positively or negatively, on the life course trajectory of an individual and his or her family. As stated earlier, QOL also incorporates health-related domains; however, in this study we focus primarily on the NHRQOL domains. We have chosen non health-related domains as a result of the neglect by researchers to incorporate these measures into previous economic evaluations of early childhood interventions. However, we must acknowledge that much of this neglect is due to methodological constraints and the reliance on the government/stakeholder perspective, which has provided results relating to savings to the criminal justice system. However, governments, non-government funding bodies

and individuals and their families are becoming more interested in understanding the potential salient benefits resulting from early childhood interventions across the life course (Nagin, 2001).

#### **4.2.1: Domains of Quality of Life**

QOL must be viewed on a number of levels. The three level model of QOL (Spilker, 1996b) provides a generally accepted approach to identifying a basic set of levels that constitute the scope of QOL (Figure 4.1); although it must be acknowledged that variations may occur regarding the exact number and definitions of levels. The top level of the pyramid describes an individual's satisfaction with life and personal wellbeing. The middle section represents the number and identity of QOL domains. The bottom level includes all the components of each domain that are assessed by the QOL tests (e.g. components of the academic skills domain may include initiative and readiness to learn).

**Figure 4.1: Levels of quality of life**

Adapted from: Spilker, B. (1996a). Introduction. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd ed., pp. 1-10). Philadelphia, New York: Lippincott-Raven Publishers.

#### **4.2.2: Non-health related quality of life (NHRQOL)**

Much of the literature on quality of life discusses health related outcomes. Unfortunately, the literature regarding non health-related outcomes is limited. Consequently, our discussion on outcome domains associated with non health-related quality of life is restricted to Spilker (1996a). Four domains are identified that impact upon NHRQOL: (a) personal-internal, (b) personal-social, (c) external-natural environment, and (d) external-societal environment (Spilker, 1996b). Each domain consists of a number of components. For example, the personal-internal domain may include components such as values and beliefs, coping strategies and spiritual status. Components such as social networks, family structure and educational status constitute components of the personal-social domain (see Table 4.1). However, it

must be acknowledged that there may be an interaction between the different components. This is discussed in detail in Section 4.3.

**Table 4.1: Domains and components of non-health related quality of life**

<b>Personal-internal</b>	<b>Personal-social</b>	<b>External-natural environment</b>	<b>External-societal environment</b>
<ul style="list-style-type: none"> <li>-Values and beliefs</li> <li>-Desires and goals</li> <li>-Personality attributes (e.g. motivation, perceived control, self-sufficiency)</li> <li>-Coping strategies</li> <li>-Spiritual status</li> </ul>	<ul style="list-style-type: none"> <li>-Social networks</li> <li>-Family structure</li> <li>-Family Functioning</li> <li>-Social Groups</li> <li>-Financial status</li> <li>-Vocational status</li> <li>-Educational status</li> <li>-Behavioural status</li> <li>-Social responsibility</li> </ul>	<ul style="list-style-type: none"> <li>-Land (pollution)</li> <li>-Quality of air</li> <li>-Quality of water</li> </ul>	<ul style="list-style-type: none"> <li>-Cultural institutions and opportunities</li> <li>-Religious institutions and opportunities</li> <li>-Shopping facilities and opportunities</li> <li>-Medical facilities and services</li> <li>-Government and political policies, institutions</li> <li>-Personal safety in the environment</li> <li>-Transportation and communication systems</li> <li>-Social and recreational institutions and facilities</li> <li>-Community spirit and demographics</li> <li>-Business institutions</li> </ul>

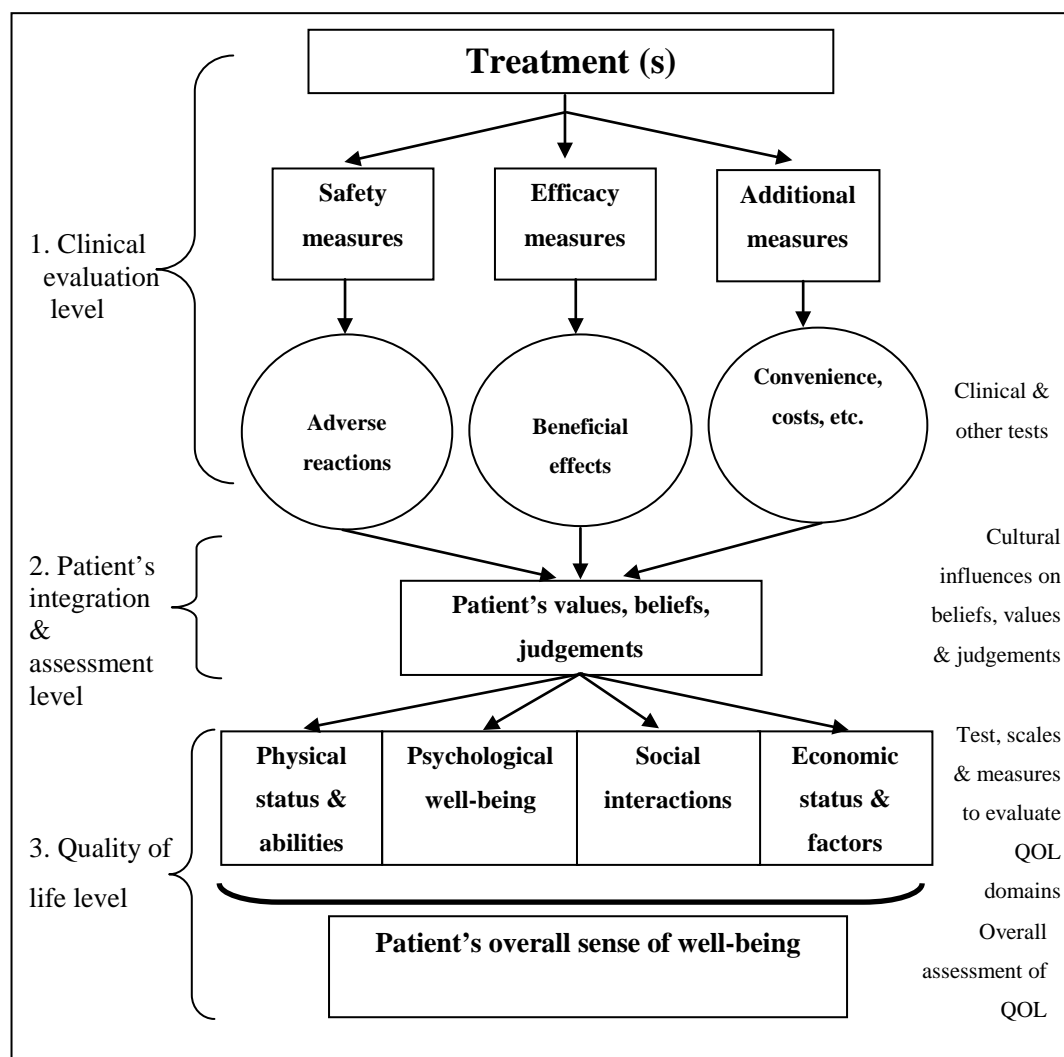
Adapted from: Spilker, B., & Revicki, D. A. (1996). Taxonomy of Quality of Life. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 25-40). Philadelphia, New York: Lippincott-Raven Publishers.

Factors both internal and external may affect an individual's perception of functioning and overall wellbeing. Internal-social factors such as social networks, including friends and family, values and beliefs, desires and goals and coping strategies may improve the psychological wellbeing of an individual who faces

limited social resources (Hommel, 2005; Hommel et al., 2006; The Developmental Crime Prevention Consortium, 1999)(see Table 3.1).

External factors contribute positively to, or negatively affect NHRQL. For example, factors in the natural environment such as air, water and land quality contribute either positively or negatively to NHRQL. An individual's external-societal environment will also influence NHRQL; including cultural institutions and opportunities, schools, medical facilities, government and political policies and personal safety (Table 4.1) (Spilker & Revicki, 1996).

A corollary of Spilker's model (Figure 4.2) is that conclusions regarding a patient's improvement in, or deterioration of, QOL are beyond simply measuring adverse reactions to, or clinical benefits of, an intervention or treatment. Ideally, it is necessary to measure all or most QOL domains for a patient or group of patients. However, in reality, this will not always be possible. Therefore, the goal should be to capture all possible information within the domains being researched. The important point to remember here is ensuring that adverse reactions to, or benefits of, the treatment/intervention are filtered through the patient's/individual's values, beliefs and judgments (Spilker, 1996b). Although Spilker's model is primarily focused on health-related QOL, the model can be easily adapted to NHRQOL domains.

**Figure 4.2: Measuring quality of life using the Spilker model**

**Source:** Spilker, B. (1996a). Introduction. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd ed., pp. 6). Philadelphia, New York: Lippincott-Raven Publishers.

#### 4.3: Measuring the domains of NHRQOL and mapping their interactions

Traditional QOL instruments are associated with personal health aspects. Currently, adequate instruments do not exist to measure non health-related domains of QOL. This however does not mean that domains of NHRQOL cannot be assessed using current instruments (Spilker & Revicki, 1996). Rather, NHRQOL domains can be assessed by adapting current health-related instruments and reviewing the longitudinal research to determine domains of interest, in particular, what domains

and items are predominantly researched in the crime prevention/developmental literature, how the domains are affected (e.g. interaction between domains), and what tools are used to measure the indicators under each domain (e.g. psychometric test tools).

Before beginning to develop a NHRQOL instrument, one should be well acquainted with the interaction between NHRQOL domains, and the various levels of aggregation of QOL. Moreover, an understanding of the categories of instruments to measure QOL is essential. This allows for a more streamlined and comprehensive development of an instrument to measure NHRQOL in the area of choice. Non health-related domains (personal-internal, personal-social, external-natural environment, external-societal environment) may interact in various ways resulting, due to their interaction, in an impact on an individual's NHRQOL. Some influences are directly related to an individual's life events and these may occur in any of the four domains.

Table 4.2 provides examples of life events. However, "most life events that occur during a clinical trial are independent of both the treatment received and the process involved" (Spilker, 1996a, p.468). Nevertheless, a direct or indirect connection exists (Spilker, 1996a). For example, a child who is provided more opportunities due to the positive impact of a preschool intervention program or a family independence program, may perform better in school, display less problem behaviour, graduate from high school, and find a job. However, we must acknowledge that the opposite may occur. Therefore, both positive and negative life experiences may result from, either directly or indirectly, participation in an intervention program.



Consequently, the relationship of all events, whether positive or negative must be assessed (Spilker, 1996a).

**Table 4.2: Example of possible life events**

<b>A. Family</b>	<ol style="list-style-type: none"> <li>1. Pregnancy, low-birth weight, abortion</li> <li>2. Marriage, separation, divorce</li> <li>3. Family disruption (e.g. quarrels and fights)</li> <li>4. Death of relative or friend</li> <li>5. Illness</li> <li>6. Important changes in children's lives that influence their parents</li> </ol>
<b>B. Work</b>	<ol style="list-style-type: none"> <li>1. New job, new responsibilities</li> <li>2. Job instability</li> <li>3. Low paid job with poor conditions</li> <li>4. Quarrels with supervisors/co-workers</li> </ol>
<b>C. Social</b>	<ol style="list-style-type: none"> <li>1. Personal habits altered</li> <li>2. Social interactions change</li> <li>3. Eating, sleeping and other patterns altered</li> <li>4. Change in time spent in recreation, hobbies, and sports</li> <li>5. Living conditions altered</li> </ol>
<b>D. Economic</b>	<ol style="list-style-type: none"> <li>1. New loans taken out and repaid</li> <li>2. Self-supporting (is about support ones-self financially)</li> <li>3. change in economic status</li> <li>4. Major purchases made</li> </ol>

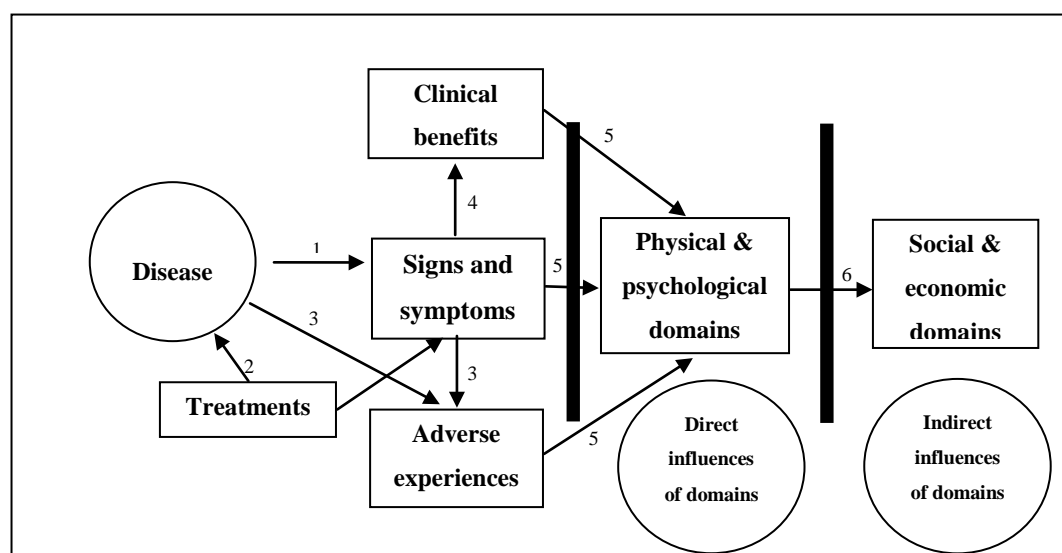
---

Adapted from: Spilker, B. (1996a). Evaluating and comparing life events data with quality of life data. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd Edition ed., pp. 468). Philadelphia: Lippincott-Raven Publishers.

Figure 4.3 provides another illustration of interaction between factors which influence the various quality of life domains. In this example, provided by Spilker (1996, p.28), the condition leads to a collection of clinical signs and symptoms (1). A treatment is then directed at the underlying problem/disease and/or eliminating signs and symptoms (2). Treatment may result in adverse experiences resulting from attempts to treat the underlying problem or the signs and symptoms resulting from the problem/disease (3). Treatment directed towards the problem/disease may result in

benefits (both health and non-health related) (4). Signs and symptoms, benefits and adverse outcomes filter through the patient's beliefs, values and judgments, which influence both physical and psychological domains (5). Influences on these domains (physical and psychological) may in turn pass through the patient's beliefs, values and judgments to influence other domains (e.g. social and economic) (6). For example, an individual who feels better both physically and psychologically may develop better social skills and improve his/her academic or work performance.

**Figure 4.3: Factors which may influence the various quality of life domains**



Source: Spilker, B., & Revicki, D. A. (1996). Taxonomy of Quality of Life. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 29). Philadelphia, New York: Lippincott-Raven Publishers.

The domains which are pertinent for HRQOL and NHRQOL are relevant for both the individual and the entire community. As such, QOL can be dealt with from an individual, community, national or global perspective. Domains such as collective physical, psychological, social, economic and spiritual can be further expanded when larger levels of aggregation are studied (Spilker & Revicki, 1996).

Levels of aggregation can be divided into six basic categories. These include the home, community, region, nation, group of nations, and the world. The six levels may be combined or further expanded depending whether there is a blending from one level to the next. The specified description of QOL for the various levels of aggregation will vary in some cases, but the terms used to quantify and illustrate findings will be similar (Spilker & Revicki, 1996). Moreover, expanding from a lower category to a higher level category (e.g. the individual and home to the community level) may be improved by conceptualising the levels of QOL which are considered pertinent to the individual or the disease/problem. However, moving from a lower level to a higher level is ultimately dependent upon on the needs of the researcher or the circumstances of the situation (Berg, Hallauer, & Berk, 1976; Spilker & Revicki, 1996). The important point to note is that the levels of aggregation may be combined or further divided into less or more categories depending on the individual needs of the researcher.

Instruments exist to measure both health- and non health-related aspects of quality of life. However, the majority are health related. Some instruments are better than others, with those that are most frequently used and well validated being referred to as “core instruments” (Spilker, 1996b).

Spilker (1996a) uses 5 categories to broadly define instruments that measure QOL. These include instruments that:

1. centre on parameters universally agreed to be elements of QOL for all the population (health-related);

2. focus on parameters that many consider to be elements of QOL for the population and are primarily used to evaluate QOL (health-related);

3. centre on parameters most believe to be part of QOL for individuals with a specific disease/symptom or particular characteristic (e.g. elderly) and are primarily used to evaluate QOL (health-related);

4. focus on parameters that are usually seen as clinical measures such as depression scales, pain scales, and tests of cognitive function; and,

5. focus on real issues to QOL but may be used to assess a component of quality of life (e.g. social functioning, social responsibility).

Spilker (1996a) argues that the above five categories move progressively outward from the core group of instruments to peripheral groupings, similar to that of concentric rings.

#### **4.4: Instruments used to operationalise health and non-health related QOL domains**

Measuring health and non-health related quality of life involves the development of an instrument that, when coupled with an economic method for attaining preference values, will evolve into a set of utility weights, or preference values. As stated in Chapter 1, the identification and measurement of non health-related outcomes associated with early childhood intervention will allow researchers and policy-makers to develop a structured process for making complex decisions under uncertainty with respect to policy options of early childhood intervention/crime prevention programs. Further, to develop a method that identifies a common metric or

set of common metric outcomes that accurately value the qualitative improvements in an individual's life from a cost-utility perspective. In other words, salient outcome domains/indicators (adolescence in this case), which are derived from longitudinal research and psychometric test instruments and their perceived preference values may be used to measure the economic utility of individual or groups of programs/interventions. Moreover, policy makers will have a structured method for making complex decisions which incorporate all relevant components in the decision-making framework. Greatest interest is in the development of an instrument that will produce a set of values that are universally transferable, in other words have a common metric, across programs within the given field. To date, considerable work has gone into developing methods for capturing health-related quality of life values. However, little work has focused on capturing non-health related quality of life values, particularly in the area of crime prevention/early intervention.

#### **4.5: Instrument to gauge improvements in NHRQOL during adolescence resulting from early childhood intervention programs**

Juniper, Guyatt and Jaeschke (1996) identified a series of steps in the development and validation of a new quality of life measurement instrument. This method has been adapted to suit the current study. The generic steps are as indicated in Table 4.3.

**Table 4.3: Generic steps in the development and validation of a new non health-related quality of life instrument**

<b>A. Development</b>	<b>B. Testing</b>
1. specifying measurement goals	1. Pretesting
2. Item generation	2. Reliability
3. Item reduction and refinement	3. Measuring consistency
4. Questionnaire formatting	4. Conduct sensitivity analysis

#### **4.5.1: Specifying measurement goals**

Defining what the instrument is to measure will aid in designing appropriate development and testing procedures. Moreover, it helps others identify whether the tool is appropriate for use within the population of interest (Bergner, Bobbitt, Carter, & Gilson, 1981; Juniper et al., 1996). Specification of measurement goals must take into account: the affected population (exclusion and inclusion criteria e.g. age, literacy level, socio-economic characteristics); the instrument's primary purpose (e.g. is the instrument evaluative, discriminative, or predictive) (see Table 4.4); the function of the instrument (what areas associated with condition are to be included e.g. physical, educational, social/emotional, occupational, deviance or criminological), and the instrument's format (e.g. interviewer and/or self-administered, and approximate number of items the instrument will contain) (Juniper et al., 1996).

The design and development of the NHRQOL instrument to be used in this thesis aims to measure the impact of various alternative forms of early childhood intervention (e.g. structured preschool program, centre-based childcare/developmental day care, home visitation, family support services and parental education) on the

development or enhancement of child, parent-child, and family NHRQOL during the adolescent life years.

**Table 4.4: Instrument development and validation in quality of life measurement**

	<b>Discriminative</b>	<b>Predictive</b>	<b>Evaluative</b>
Item generation	Identify all items of impairment that might be important	Identify all items of impairment that might be important	Identify all items of impairment that might be important
Item reduction	Delete items common to all individuals	Delete items common to all individuals	Select the most frequent and important items
Response options	Response options adequate to achieve fine or coarse discrimination, depending on goals	Response options adequate to predict criterion standard	Response options with sufficient gradations to register within-individual change
Reliability	Large and stable inter-individual variation	Large and stable inter-individual variation	Not relevant
Consistency	Not relevant	Not relevant	Able to detect small changes in consistency over time
Validity	Cross-sectional construct validity	Cross-sectional criterion validity	Longitudinal construct validity

Adapted from: Juniper, E. F., Guyatt, G. H., & Jaeschke, R. (1996). How to develop and validate a new health-related quality of life instrument. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 50). Philadelphia, New York: Lippincott-Raven.

Five alternative forms of early childhood intervention programs are included in this thesis: (a) Structured preschool program (SPP) (e.g. 5 x half days for 3-4 year olds) (b) Family/parenting support including education, guidance, case management, and referrals to other agencies (FSS) (e.g. health and other human services), (c) Centre-based childcare/developmental day care (CBCC), (d) Home visitation programs (HV), and (e) Parental education (PE) (groups or classes exclusively for parent's, e.g., group parent or individual parent training).

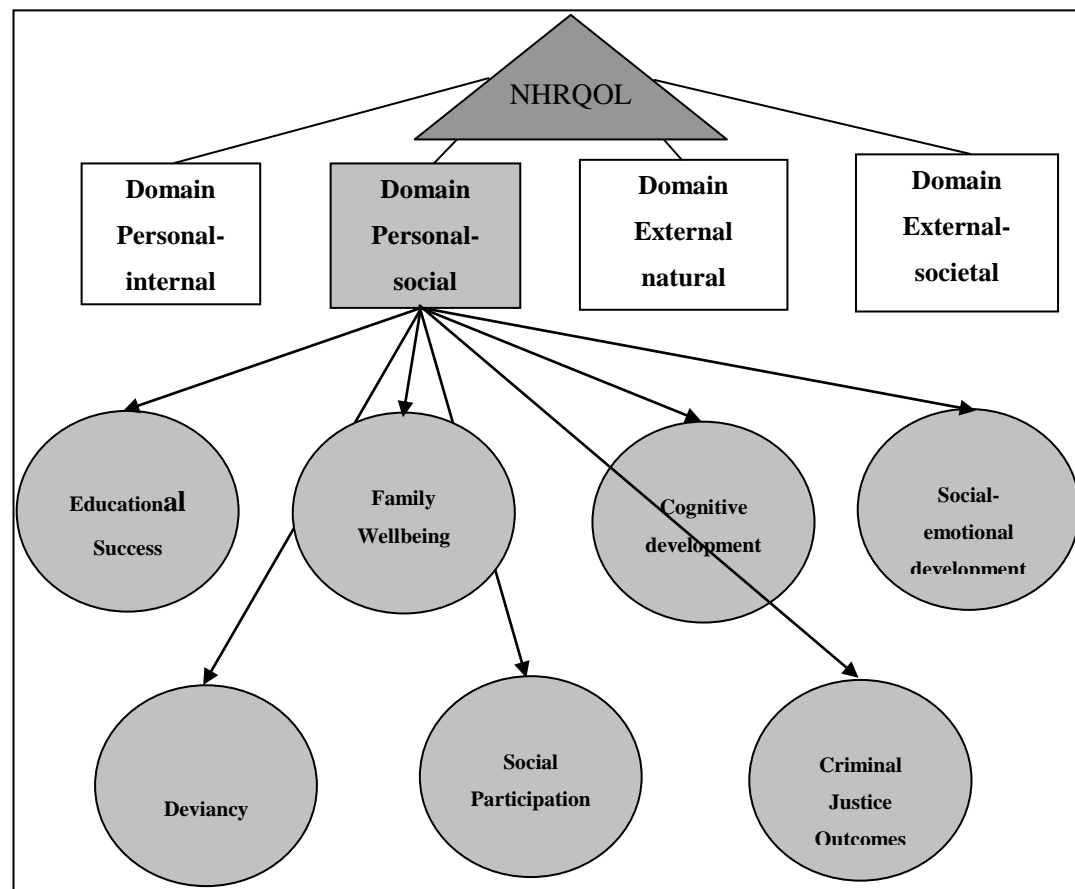
We have chosen these categories of program given the depth of outcome data relating to comparable programs available from longitudinal research focusing on adolescent life follow-ups. Unfortunately, programs such as parent participation groups in school settings, and playgroups and pre-kindergarten programs have not been sufficiently analysed during the adolescent life phase. Consequently, it was not possible to gauge the potential effect on NHRQOL of such programs.

The primary purpose of the instrument is evaluative (Table 4.4). Although there are four broad domains of NHRQOL possible for analysis, this thesis will focus exclusively on the personal-social domain (Table 4.1, Column 2). This domain has been selected due to the emerging interest by policy-makers and researchers in crime prevention/early intervention field on the economic evaluation of the salient benefits to the individual that arise from early intervention programs. This domain is further disaggregated into components relating to educational success (ES), cognitive development (CD), social-emotional development (SED), deviancy (D), social participation (SP), criminal justice outcomes (CJ) and family wellbeing (FW). (See Figure 4.4).

#### **4.5.2: Item generation**

To develop an instrument which incorporates items that are specific to the condition of interest (in this case outcomes associated with early childhood programs during the adolescent life phase) one can conduct unstructured interviews, hold focus groups, review the specific literature, and/or review other alternative instruments highlighting relevant items (Gold, Patrick, Torrance, Fryback, Hadorn, Kamlet, Daniels, & Weinstein, 1996a).



**Figure 4.4: Domains and individual components in the current instrument**

**Note:** Domains are indicated by rectangles while components of a particular domain are indicated by circles.

In this thesis, a meta-analysis of longitudinal research relating to the adolescent outcomes associated with early childhood intervention programs is employed to identify the relevant items for inclusion in our instrument. A meta-analytic approach was chosen since it permits a systematic identification of the relative effectiveness of programs on multiple dependent variables, both in the short- and long-term (Glass, McGaw & Smith, 1981; Lipsey & Wilson, 2000). A detailed outline of the meta-analytic approach is provided in Chapter 5. In short, the meta-analytic technique provides a tool for combining expert opinions on what items should be included into a decision-making framework for evaluating programs from a cost-utility perspective. In doing so it provides a vehicle for bringing together various

studies from a diverse group of researchers and clinicians from different countries at alternate points in time. This, therefore, provides a highly enriched longitudinal perspective on the impact of early childhood intervention programs on the salient outcomes during adolescence.

A thorough investigation has also been conducted relating to the psychometric tools that can be used to assess the effectiveness of early childhood interventions on outcomes during the life course, including the adolescent life phase. The results of this investigation are provided in Appendix A. This investigation enabled the identification of relevant indicators to focus on when organising our meta-analytic methodology.

#### **4.5.3: Item reduction and refinement**

Once a pool of components and indicators has been generated (within the given domain), the investigator must then choose the most suitable items for inclusion in the final instrument (Juniper et al., 1996). Two approaches are available: one may ask experts (e.g. doctors, clinicians, patients or others directly involved in, or affected by the condition of interest) to identify, from a list, items they have experienced or have been affected by as a result of their condition/illness/experience etc. (usually adopted in measuring HRQOL) (Juniper et al., 1996); or items may be selected based on outcomes identified in a literature review or meta-analysis. The latter approach is adopted in this thesis. Normally, items chosen to report on by longitudinal researchers represent the most salient outcome variables for the target population. Often, their choices are influenced by items reported on by previous researchers/clinicians, but

funding bodies and those directly involved in the intervention/program can also influence their choice.

Moreover, longitudinal research tells us what can be changed and by how much (Maxfield & Babbie, 2001). Mathematical modelling techniques such as factor analyses are adopted by some investigators to determine items for inclusion and exclusion. Items that have strong correlations are grouped together; whilst items that do not have a strong association with one of the key factors are excluded from the final questionnaire (Maxfield & Babbie, 2001). One important consideration when reducing items is to ensure that items selected represent all sub-groups within the population affected (Juniper et al., 1996; Spilker, 1996b).

The drawback of the factor analytic approach, as highlighted by Juniper et al. (1996), is that if this analysis arrives at counter-intuitive groupings, difficulty arises in how to proceed. Therefore, using intuition and clinical experience to decide on what items belong in the selected domain or domains appears to be the most appropriate method. If uncertainty exists about some items, then correlations of items with other closely related items within a given domain can be used (Juniper et al., 1996).

#### **4.5.4: Questionnaire formatting**

The development of an evaluative instrument requires that one carefully select response options. Response options, in this example, refers to the categories or scales that are available for responding to the questionnaire options (Bergner et al., 1981; Juniper et al., 1996). One must ensure that the instrument is responsive to important changes regardless of their size. To ensure that this measurement property is

enhanced, investigators often choose scales with a number of options. For example, a 7-point scale where responses may range from 1 to 7: one representing a very low achievement in relation to a particular outcome, and seven a very high level of achievement. Health economists tend to adopt the use of either a Likert scale or a visual analogue scale for obtaining responses to questionnaire options. Both approaches yield similar data, however, the Likert scale tends to have the advantage of being both easier to administer and interpret (Drummond et al., 1997).

Secondly, a time specification should be included. For example, respondents may be asked how they rate achievement related to a particular outcome over a defined period of time. The important thing to note here is that there is a respondent's upper limit of accuracy to recall. Time specifications clearly need to be tailored to suit the study as the time horizon differs more significantly from the current point of time.

The next item concerns questionnaire administration. Juniper et al., (1996) argue that if a questionnaire is given more than once to a respondent, previous responses should be shown as it improves the validity of the questionnaire. Finally, language used in the questionnaire should avoid jargon, idioms, or metaphors. This enhances its usefulness across countries and cultures.

The response option used in the instrument developed in this thesis will be adapted from a methodology outlined by Saaty (1990) and Isard and Smith (1982). In short, the respondent judges the importance of each criterion in pair-wise comparisons, which are expressed on a semantic scale of 1 (criterion are judged

equally important) to 9 (i.e. one criteria is judged to be 9 times more important than the one to which it is being compared). This method requires a respondent to judge the relative importance of a set of pair-wise comparisons (expressed by posing the questions “which of the indicators is more important?” and secondly, “by how much?”) based on subjective evaluations drawn from the respondent’s cumulative life experience to date.

#### **4.4.5: Instrument testing**

When a new questionnaire is first developed and administered, it must be pre-tested to ensure all questions, including wording and meaning are clearly and correctly understood (Bergner et al., 1981; Juniper et al., 1996). In this thesis, the questionnaire was administered first to five respondents from various backgrounds (e.g. clinicians, preschool staff, academic researchers and policy-makers). Pre-test respondents were selected to represent as wide a spectrum as possible relating to policy decision-makers associated with early childhood interventions. After an uninterrupted administration of the questionnaire, respondents were asked to explain in their own words their understanding of each item. Consistency problems regarding wording and understanding were noted, and necessary changes were made to the draft questionnaire. A revised questionnaire was then administered to five new pre-test respondents. The process continued until no more changes to the questionnaire were indicated as necessary to address concerns raised by pre-test respondents.

#### **4.5.6: Reliability**

To ensure reliability of the instrument, it is important to keep all other sources of “noise” to a minimum. For example, using one interviewer, keeping conditions the same across all interviews (e.g. time of day, quiet and distraction free environment), and ensuring that participants selected to take part in the testing of the instrument represent the population in which the instrument is to be used (Juniper et al., 1996). The application of the instrument in this thesis employed the approaches described above in an effort to ensure that “noise” was kept to a minimum.

#### **4.5.7: Consistency**

Consistency is essential in any evaluative instrument. A good instrument will be able to detect changes in consistency and the analyst should administer appropriate protocols for exclusion of unacceptably inconsistent questionnaire responses. A number of approaches are available to test for consistency, however, in this thesis we will adopt the method developed by Saaty (1990; 2000), which calculates a consistency ratio (C.R.) and recommends guidelines for non-inclusion or inconsistent responses. This approach is discussed in more detail in Chapter 8.

#### **4.5.8: Sensitivity analysis**

The stability of results should be tested using a simple sensitivity analysis to help measure the responsiveness of the results to slight changes in judgment values

(Alexander & Saaty, 1977a; Johannessen, Bandara, & Smith, 2004; Saaty, 1990). Saaty (1990) identifies three ways to test sensitivity: (1) by deriving a mathematical estimate of the fluctuation (measure of consistency), (2) by obtaining answers based on a large number of computer runs designed to test the sensitivity, and, (3) by using a combination of methods 1 and 2. In this thesis, method 3 is adopted.

#### **4.5: Conclusion**

In this chapter we have discussed the broad concepts of quality of life (HRQOL) and non health-related quality of life (NHRQOL). Additionally, we sketched briefly the methodology used in this thesis to classify non health-related quality of life outcomes domains and their respective indicators associated with early childhood intervention during adolescence. Moreover, we have identified and discussed a series of steps to be employed in the development and validation of a new quality of life measurement instrument. These steps include the development of the instrument (e.g. specifying measurement goals, item generation, item reduction and refinement and questionnaire formatting), and testing of the instrument (e.g. pre-testing, reliability, measuring consistency and conducting a sensitivity analysis).

Chapter 5 of this thesis provides the methodology of the meta-analytic technique applied to estimate the effectiveness (in terms of effect size) of early childhood intervention programs on outcomes situated within the adolescent life phase. Further, in Chapter 6 we provide the results of our meta-analysis, and discuss the effects of 5 forms of early childhood intervention (structured preschool program (SPP), home visitation (HV), centre-based childcare/developmental day care (CBCC),

family support services (FSS) and parental education (PE)) on at-risk children and their families on seven outcome domains (educational success, cognitive development, social-emotion development, deviancy, social participation, criminal justice outcomes, and family wellbeing) during adolescence (age 12-19 years).



## **CHAPTER 5: AN ANALYSIS OF LONGITUDINAL RESEARCH ON EARLY CHILDHOOD INTERVENTION PROGRAMS**

### **5.1: Introduction**

This chapter provides a detailed description of the meta-analytic methodology used in this thesis, which identifies best estimates of the effectiveness (in terms of effect size) of early childhood intervention programs on outcomes during adolescence based on published evaluations of such programs. It also provides a detailed overview of outcome domains (non health-related) considered significant during adolescence. Section 5.1 of the chapter provides a brief overview of meta-analytic methodology in general, including a discussion of benefits and criticisms of this approach. Moreover, it incorporates the calculation of individual effect sizes and variances, calculation of weighted mean effect size and other statistical applications pertinent to the meta-analytic method. Further, issues associated with the interpretation of meta-analytic results are discussed. Section 5.2 discusses previous meta-analyses relating to early childhood intervention programs. Section 5.3 provides a detailed description of the particular meta-analytic approach adopted in this thesis including the meta-analytic framework, criteria for selection, the unit of analysis, types of early childhood programs analysed, coding of variables, outcomes domains analysed and the calculation of effect sizes. Finally, Section 5.4 provides a summary of the chapter and briefly introduces Chapter 6.

#### **5.1.1: Overview of the meta-analytic methodology**

Traditionally, meta-analyses have been conducted to estimate the relative effectiveness of intervention programs on a series of dependent variables, in both the

short and long-term. A meta-analysis is a statistical method for combining results from a series of studies, whereby researchers are able to correct for various statistical artefacts and to aggregate results in order to obtain an estimate of the relationship between variables, or address a series of related research hypotheses (Durlak & Lipsey, 1991; Lipsey & Wilson, 2000). The benefit of the meta-analytic technique is that the method can accurately summarise existing knowledge relating to the relative effectiveness of prevention programs on multiple dependent variables, both in the short- and long-term (Glass et al., 1981; Lipsey & Wilson, 2000). Moreover, the meta-analytic technique allows researchers to manage large amounts of potentially salient information for later analysis.

For the purpose of this research, this methodology has been adapted to serve two functions. First, the methodology is used to identify the salient outcome domains and their indicators during the adolescent life phase. Secondly, it is used to estimate the effectiveness of early childhood intervention programs in terms of the effect size on a range of outcome domains. Information on outcome domains, indicators and effect size provides the basis for developing a hierarchical decision-making structure for use in Chapters 8 and 9 for selecting between alternative early childhood intervention programs.

Currently, longitudinal studies highlight the effectiveness of various forms of early childhood interventions across 5 life phases: early childhood (age 0-5 years, including preschool); childhood (age 5-12 years, including transition to primary school); adolescence (age 12-19 years, including transition to high school and the transition from high school to the work force); early adulthood (age 20- 27 years);

and, adulthood (28+ years). As indicated previously, this thesis focuses on outcomes associated with the adolescent years (age 12-19 years) with particular attention on outcomes surrounding the transitions into and out of this life phase (transition from primary school to high school and high school to the workforce or post secondary education).

What sets this meta-analysis apart from past analyses is that first, it analyses a series of dimensions never before measured. Previous meta-analyses of early childhood intervention programs (Nelson et al., 2003) have focused on three broad outcome constructs, namely cognitive development, social-emotional development and parent/family well-being. This analysis reanalyses those broad outcome domains in more detail and includes more up-to-date results from current longitudinal research. It also adds further dimensions by analysing levels of academic/education success (e.g. school/academic success, reduction in special education), deviance, social participation, and effects on involvement in the criminal justice system in the adolescent years. Although a meta-analysis was conducted by Farrington and Welsh (2003) reviewing the effectiveness of family-based prevention programs in reducing offending and antisocial behaviour by children and adolescents, there are significant differences between this study and that by Farrington and Welsh. First, Farrington and Welsh's meta-analysis focused primarily on short-term effects (typically, 1 year post intervention). Secondly, Farrington and Welsh's focal point on long-term effects (typically up to 3 years post intervention) was not focused exclusively on the adolescent life phase. Rather, it included a large proportion of outcomes that related to school-based programs and programs with older children. Moreover, not all programs included in the analysis by Farrington and Welsh focused exclusively on at-risk

communities and their children. Finally, the present meta-analysis is different from other meta-analyses in that rather than calculating a mean effect size (d.) alone, ranges of effect sizes for dependent variables are presented. This issue is discussed in greater detail later in this chapter.

### **5.1.2: Benefits of meta-analysis**

The benefits of adopting a meta-analytic approach are that it allows estimation of the relative effectiveness of prevention programs on multiple dependent variables, both in the short- and long-term (Glass et al., 1981; Lipsey & Wilson, 2000). The method allows the researcher to keep track of large amounts of potentially important information for use in later analysis. This quantitative coding of key results from diverse prior studies makes explicit the sample of studies selected for possible future replication. For the purpose of this thesis, the meta-analytic results (relating to knowledge accumulated worldwide in relation to the effectiveness of programs adopted in multiple settings in terms of geographic locations, ethnic groups and other moderating influences) can be made available to policy makers and survey participants to view in summary form (Durlak & Lipsey, 1991; Lipsey & Wilson, 2000; Wolf, 1986).

### **5.1.3: Criticisms of meta-analysis**

Since its inception (Smith & Glass, 1977), meta-analysis has been heavily criticised. Several threats to the validity of its results have been raised, with the three most prominent areas of criticism highlighted here. The first area of criticism has been labelled “apples and oranges” (Sharpe, 1997). Criticisms falling in this category (e.g.

Eysenck, 1978; Rachman & Wilson, 1980) argue that meta-analysis combines statistical results from studies that “measured different things, manipulated different variables, and tested different subject populations” (Sharpe, 1997, p.882). However, a careful system for inclusion and exclusion criteria can help minimise this threat to validity. The second area of concern relating to validity has been labelled “file drawer”. Rosenthal (1979) for example argues that problems can arise from a failure to obtain a representative sample of the population of studies relating to the topic. That is, publication bias is potentially increased due to the tendency of journals to publish studies reporting statistically significant findings, and a reduced tendency to publish studies with statistically non-significant results (Lipsey & Wilson, 2000). However, Wortman (1994) recommends that researchers calculate a “fail-safe  $n$ ” (Rosenthal, 1994) to estimate the number of “null hypothesis accepting” studies that would be required to reduce the overall effect size to a non-significant level. The third area of concern relating to validity has been labelled “garbage in, garbage out” (Eysenck, 1978). Rachman and Wilson (1980) argue that Glass and colleagues purposely designed the methodology in such a way as to promote inclusion of all prior studies without the use of academic judgement because “...their fear of selection bias was greater than their fear of including studies of low quality... [and] combining methodologically poor studies led to a distorted picture of psychotherapy effectiveness” (Sharpe, 1997, p.882). However, once again, a carefully designed methodological structure, which incorporates criteria for inclusion and exclusion, can help to reduce this problem.

Other proposed weaknesses of the method include: the effort and expertise required to conduct a thorough meta-analysis; its structured and somewhat mechanical

approach; and, the most persistent criticism, the mix of studies included (Durlak & Lipsey, 1991; Lipsey & Wilson, 2000). Lipsey and Wilson (2000) observe; “Critics argue, with some justification, that mean effects sizes and other such summary statistics produced by meta-analysis are not meaningful if they are aggregated over incommensurable study findings” (p.8).

#### **5.1.4: Effect size protocols used in meta-analysis**

Meta-analysis from a statistical standpoint involves the following steps: (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 ( $\underline{d}$ ). The weighted mean effect size ( $\underline{d}$ ) is calculated by weighting individual effect sizes by the inverse of its variance. A general formula for calculating a weighted mean effect size is demonstrated in Lipsey and Wilson (Lipsey & Wilson, 2000, pp. 113-114); (3) calculation of 95% confidence intervals (CIs) surrounding those effect sizes; (4) calculation of median effect sizes; and (5) calculation of the Q statistic to assess the heterogeneity of the effect sizes for each predictor category. The Q statistic is used to evaluate the heterogeneity of individual effect sizes contributing to the weighted  $\underline{d}$ ; and (6) re-calculation of weighted mean effect sizes and 95% CIs based upon the extent of heterogeneity in the effect size distributions.

#### **5.1.5: Calculation of individual effect sizes and variances**

To determine the relative magnitude of the experimental treatment, in other words, the size of the experimental effect, Cohen’s  $d$  is used. The advantage of using Cohen’s  $d$  is firstly its burgeoning popularity, thus making it a standard method of

determining effect-size measurements. Calculating  $d$  allows immediate comparison of an increasingly growing number of published research studies (Durlak & Lipsey, 1991). Moreover, based on Cohen's (1992) proposition that effect sizes of .20 are small, .50 are medium, and .80 are large, one can compare effect sizes to established benchmarks. However, it must be acknowledged that Cohen's descriptions of effect sizes (e.g. .20 = small) may be misleading. Given that ranges of effect sizes in crime prevention literature often describe an effect of .20-.40 to be a reasonably good effect on the population concerned, one must be careful not to equate an effect of .20 in terms of a small and insignificant effect; but rather, a positive effect in terms of outcome but a small effect in terms of scale comparison (e.g. .20 compared to .85).

When means and standard deviations are reported,  $d$  is calculated by subtracting the mean score ( $X_2$ ) (post-test follow-up) for the comparison group from the mean score of the intervention group ( $X_1$ ) and dividing the result by the pooled standard deviation ( $S_{pooled}$ ), where  $N_1$  is the number of participants in the intervention group and  $N_2$  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 (see Equation 5.1)

$$ES_{sm} = \frac{\overline{X_2} - \overline{X_1}}{S_{pooled}} \quad (5.1)$$

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

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  $X^2$  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), where  $n$  is the total sample size and  $ES_{sm}$  is the biased standardised mean difference (see Equation 5.2).

$$ES_{sm} = \left[ 1 - \frac{3}{4n - 9} \right] ES_{sm}' \quad (5.2)$$

The variance (Var) of  $d$  was then calculated using Equation 5.3, where  $(n)$  are the sample sizes for each group and  $(d)$  is the mean:

$$Var\ d = \left[ \frac{n_1 + n_2}{n_1 n_2} + \frac{d^2}{2(n_1 + n_2)} \right] \quad (5.3)$$

#### 5.1.6: Calculation of weighted mean effect sizes for each predictor variable

Following the calculation of individual effect sizes and variances (extracted from the primary studies), a weighted mean effect size ( $\underline{d}$ ) can be calculated using Hedges and Olkin's (1985) formula, where  $\underline{d}$  is the weighted mean effect size,  $k$  is the number of findings,  $w_i = 1/v_i$ , and  $v_i$  is the variance of the individual effect size (see Equation 5.4). This involves weighting each individual effect size by the inverse of its variance.



$$\underline{d} = \frac{\left( \sum_{i=1}^k w_i d_i \right)}{\left( \sum_{i=1}^k w_i \right)} \quad (5.4)$$

Subsequently, the variance of the weighted mean effect size can be calculated using Equation 5.5:

$$Var(\underline{d}) = \frac{1}{\left( \sum_{i=1}^k w_i \right)} \quad (5.5)$$

#### 5.1.7: Calculation of 95% confidence intervals

Ninety-five percent confidence intervals are calculated to aid in interpreting the effect sizes. A wider confidence interval suggests a less reliable effect, while if the confidence interval does not contain zero this signifies a significant effect at the 0.05 level (Lipsey & Wilson, 2000). The variance of the weighted mean effect size is then used to calculate 95% confidence intervals using Equation 5.6.

$$95\% \text{ C.I.} = \underline{d} \pm 1.96\sqrt{\text{Var}(\underline{d})} \quad (5.6)$$

#### 5.1.8: Calculation of the median effect size

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, opposed to selecting one specific effect size.

#### **5.1.9: Assessing heterogeneity of effect sizes using the Q-statistic**

Next, the homogeneity of the seven outcome constructs (see Figure 4.4) can be tested. In particular the Q statistic can be used to evaluate the heterogeneity of individual effect sizes contributing to the weighted  $\underline{d}$ . Nelson, Westhues and MacLeod (2003) argue that the Q-statistic allows for a statistical evaluation of the variation in the distribution of effect sizes for a given outcome construct. Lipsey and Wilson (2000) also contend that the Q-statistic 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 found to be 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 adopted first. If Q is found to be significant, the analysis can be 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, while a random effects model assumes that the variability between effects sizes is due to sampling error and the variability in the population of effects (e.g. each study is estimating a slightly different population effect size). Thus, the fixed effects model weights each study by the inverse of the sampling variance (Lipsey & Wilson, 2000) (see Equation 5.7).

$$W_I = \frac{1}{se_i^2} \quad (5.7)$$

While, the random effects model weights each study by the inverse of the sampling variance and a constant ( $\hat{v}_\gamma$ ) that represents cross-population effects variability (see Equation 5.8).

$$W_I = \frac{1}{se_i^2 + \hat{v}_\gamma} \quad (5.8)$$

Where the random effects constant ( $\hat{v}_\gamma$ ) is calculated using Equation 5.9:

$$\hat{v}_\gamma = \frac{Q_T = k - 1}{\sum w - \left( \frac{\sum w^2}{\sum w} \right)} \quad (5.9)$$

Equation 5.10 can then be used to calculate the Q-statistic for each outcome construct.

$$Q = \sum_{i=1}^k w_i (d_i - \underline{d})^2 \quad (5.10)$$

In this formula  $k$  represents the number of findings,  $w_i$  is the inverse variance weight of the individual effect size for each finding,  $d$  is the individual effect size, and  $\underline{d}$  is the weighted mean effect size. The formula estimates the sum of the squared deviations from  $\underline{d}$ , with each finding weighted by its inverse variance, and assesses the significance of the final Q value using the chi-square distribution with  $k-1$  degrees of freedom (Nelson et al., 2003).

If the Q-statistic is found to be significant for a particular construct, the finding contributing the largest amount of variance to the overall Q-statistic can then be removed. The weighted mean effect size is then recalculated, and the Q-statistic recalculated based on the new  $\underline{d}$ . This process continues until the Q-statistic is no longer significant or the variable possesses less than two findings, in which case it can be removed from the meta-analysis.

#### 5.1.10: Fail-safe n

Often, meta-analyses are vulnerable to bias given that studies that demonstrate non-significant findings rarely get published; this is known as the fail-safe n dilemma (Hedges & Olkin, 1985). Failure to include these unpublished studies into the meta-analysis may result in biased findings due to sampling, whereby the plausibility of observed results may be called into question. To circumvent this dilemma, a method developed by Rosenthal (1979), known as the ‘fail-safe n’ formula has been 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 (Lipsey & Wilson, 2000). The method has been adapted by Orwin (1983) in order to determine the number of studies required to reduce an effect size to a specified criterion level (see Equation 5.11).

$$k_o = k \left[ \frac{\overline{ES_k}}{\overline{ES_c}} - 1 \right] \quad (5.11)$$

where  $k_0$  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 in the mean effect size, and  $ES_k$  is the weighted mean effect size (Lipsey & Wilson, 2000).

#### **5.1.11: Interpretation of Meta-Analytic Results**

Meta-analysis results may be expounded by following four guidelines: (1) effect sizes may be interpreted according to the general rules of .20 = small, .50 = medium, and .80+ = large (Cohen, 1992); (2) a confidence interval around  $d$  that does not contain zero is equivalent to the effect size being significant at  $p = 0.05$  (Hanson & Morton-Bourgon, 2004); (3) reliability of the effect size is ascertained by the connection between the weighted  $\underline{d}$  and median  $\underline{d}$  (Hanson & Morton-Bourgon, 2004; Lipsey & Wilson, 2000); and (4) reliability of the effect size is also determined on the basis of a non-significant  $Q$  statistic (Lipsey & Wilson, 2000; Nelson et al., 2003).

### **5.2: Previous meta-analyses of early intervention programs for children**

Several meta-analyses have been conducted reviewing different types of intervention programs for children (e.g. Durlak & Wells, 1997, 1998). However, as well as being relatively dated (the intervention literature reviewed in these meta-analyses being limited to those published in 1991 or earlier); the literature included did not focus exclusively on early childhood intervention programs. Rather, they also focused on secondary mental health prevention programs. Additionally, their review included analysis of programs that did not have a follow-up assessment.

More recently, MacLeod and Nelson (2000) reviewed programs designed to promote child maltreatment and family wellness. They also did not focus exclusively on early childhood programs; instead their focus was on parent-child relationships and family wellbeing. In similar fashion to Durlak and Wells, MacLeod and Nelson included studies that did not have a follow-up post intervention. A relatively recent meta-analysis (Nelson et al., 2003) did focus specifically on the effectiveness of early childhood intervention programs (what they called preschool intervention programs), however, the dependent variables in this study were limited to cognitive outcomes, social-emotional outcomes, and parent-family outcomes. Moreover, the classification by Nelson et al. of indicators under various outcome constructs was very broad, with little justification or discussion provided regarding the scientific method applied to create outcome domains or the underlying reasoning behind their subsequent groupings of outcome domains and their various indicators. All programs included the meta-analysis by Nelson et al. (2003) did have a follow-up. Although the analysis focused specifically on preschool interventions, measurements of effectiveness did not include lower level items, or relevant indicators relating to broad outcome domains (e.g. educational success – school dropout and graduation). Rather, three broad outcome domains were developed to synthesise many fine-grained outcomes into one broad outcome domain. For example, the construct ‘children’s cognitive development’ was comprised of measures of IQ, achievement tests, grades, and teacher ratings of children’s academic skill. The rationale for the use of a few broad constructs resulted from lack of data across the interventions to meta-analyse the more fine-grained constructs (Nelson et al., 2003).

### **5.3: Outline of particular meta-analytic methodology employed in this thesis**

#### **5.3.1: Aim of the analysis**

The aims of the current analysis are: to identify the multiple dependent variables over the adolescent life phase and group these variables (based on an analysis of longitudinal research) into meaningful outcome domains including their relevant indicators; and highlight the potential effectiveness of early childhood intervention programs on the seven resulting outcome domains (educational outcomes; cognitive development; social-emotional development, deviance, social participation, criminal justice, and family wellbeing) using weighted effect sizes (d.) and ranges of effect sizes.

Ranges of unweighted effect sizes (d) for a given dependent variable, together with a calculated weighted mean ES (d.) are provided. This method has been adopted so that all effect sizes are represented, rather than adjusting for outliers. In this case, outliers are extreme effect sizes that are around 2 or 3 standard deviations from the mean. Lipsey & Wilson (2000) argue that extreme effect sizes may have disproportionate influences over statistics such as means and variances, which, in turn, may distort findings. There are methods for handling outliers, for example, eliminating them from the effect size distribution or recoding extreme values to more moderate ones.

### **5.3.2: Identifying outcomes relevant to early childhood intervention across the life course**

Given that no multi-attribute classification system of non health-related outcomes of early childhood interventions exists, we set about a four month analysis of both psychometric test libraries and longitudinal research studies to identify salient outcome domains, outcomes relevant to those domains and their individual indicators. Multi-attribute systems provide a compact but comprehensive framework for describing various health and non health-related states for use in evaluation studies (Torrance, 1986). A life course perspective was adopted and relevant outcomes identified for specific life phases from early childhood to late adulthood (28 years +). The psychometric test libraries were found to be very helpful for identifying important non health-related outcomes and their indicators for the transition from preschool to primary school, and from primary school to high school. Longitudinal research studies made up the bulk of the data for the adolescent life transition (e.g. 13 – 18 years) with a few psychometric test tools (e.g. Achenbach behaviour checklist, HOME Inventory Administration Tests) utilised. Early adulthood (19-27 years of age) and later adulthood (27 years+) outcomes and their relevant indicators were identified from longitudinal research studies (e.g. Perry Preschool Program (Scheinhardt et al., 1993, Parks, 2000 .Coalition for evidence-based policy, 2005), Busselton Study (Cullen & Cullen, 1996), Abecedarian Project (Campbell et al., 2002), Chicago Longitudinal Study (Reynolds et al., 2001), Elmira Prenatal/Early Infancy Project (Olds, et al. 1997, James, 1995 ).

Data were then combined and analysed for each major life phase or transition point (e.g. transition from primary school to high school, transition from primary



school to high school, the adolescent life phase, the transition from high school to post-secondary education or the workforce, early adulthood and late adulthood). With the guidance from a group of experienced psychologists from the School of Psychology and the School of Criminology and Criminal Justice, Griffith University, salient outcome domains were developed and outcomes within the individual domains and their relevant indicators were categorised according to their relevant outcome domains.

Outcomes of this analysis are presented in Appendix A. Results presented in Appendix A represent the first comprehensive analysis and amalgamation of both psychometric test indicators and longitudinal research items for early childhood intervention program outcomes across the life course (from early childhood to late adulthood). Results highlight the salient outcome domains, sub-domains relevant to each domain and their individual indicators.

### **5.3.3: Meta-analytic framework**

The framework for this meta-analysis consists of three main components: (a) outcomes related to children's cognitive development, educational/academic success, social-emotion development, deviance and social participation, criminal justice, and family wellbeing; (b) the time period of outcome assessment (in this study the various dependent variables are studied within adolescence (age 12-19 years), which includes the transition from primary to high school and the transition from high school to work or postsecondary education); and (c) moderators of potential positive outcomes (for example, program characteristics (length, intensity of program and any supplementary program offered as follow-on to the initial preschool intervention), participant

characteristics (ethnicity and cultural background of participants), and study characteristics (location of study, methodological quality of the study)). Results are provided in Chapter 6, Table 6.2.

#### **5.3.4: Selection criteria for inclusion in meta-analysis**

The following criteria are used to select studies for inclusion in this thesis's meta-analysis:

- (a) the intervention must have begun during the child's preschool years (i.e. before the child began primary school);
- (b) the focus of the intervention must have been on developing or enhancing child, parent-child, or family well-being;
- (c) intervention programs need to have adopted either universal (e.g. all children in the population) or selected approaches (e.g. high-risk groups within the population);
- (d) intervention programs should not have been specifically aimed at treating children with mental health or severe developmental problems, as this would have biased the results;
- (e) a prospective design with control and comparison groups must have been used;
- (f) the results of research must have been reported in journal articles, books, or book chapters published in 2007 or earlier;
- (g) outcome measures that related to one or more children's cognitive development, academic/educational success, social-emotion development, deviance, social participation, criminal justice outcomes, or family wellbeing must have been

collected and reported. However, studies did not have to capture effects across all of these outcome domains;

(h) studies must have included at least one post-intervention follow-up (during the adolescent years) and collected data related to a minimum of one of the outcome domains mentioned above;

(i) studies must have been reported in such a manner that effect sizes could be identified and calculated; and

(j) studies must have been directed toward populations who live in, or come from (at the time of the intervention) disadvantaged and/or low socio-economic backgrounds.

The above criteria were chosen to ensure that data collected provided useful indicators of the effects of early childhood intervention projects on at-risk children, devoid of mental health or severe developmental problems, and their families during the adolescent life phase. Outcomes were limited to impacts of programs during adolescence. Developing eligibility criteria for rating the suitability of individual study inclusion into the meta-analysis is well supported by the literature (Durlak & Lipsey, 1991; Lipsey & Wilson, 2000), since it acts to ensure that studies chosen represent or clearly define the population of interest thereby minimising some of the methodological weaknesses identified in Section 5.2.2.

Studies included in the analysis must have met all the criteria from (a) through to (j). Some well known studies such as Healthy Families America (Daro & Harding, 1999) and Parents as Teachers (Wagner & Clayton, 1999) were not included in the analysis as a result of not meeting all the criteria. Studies that included results relating

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

The starting point in developing our meta-analytic results was a manual search of the literature through journal articles. This search focused on published articles between the years 1970 and 2007. Typical journals included in the search include: *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*, and *Journal of Prevention and Intervention in the Community*.

Additionally, a scan of relevant review articles, reference sections of articles revealed as meeting the selection criteria and asking key researchers for assistance in identifying relevant additional articles occurred. An examination of ten electronic databases covering the years 1970 to 2007 (e.g. SAGE full text, CSA, Informit) using the keywords such as prevention, early intervention, preschool education, children, home visitation, multi-component program took place. Efforts were made to track

down any unpublished studies highlighted in the search. Additionally, authors of published articles were contacted by email and other forms of communication to ascertain if additional results relating to our key outcome domains were available. This has significantly contributed to our meta-analytic results since previous meta-analyses have been unable to include evaluations for some programs due to lack of relevant information. Fortunately, some of this missing data has been able to be tracked down as a result of this personal approach to key researchers. Over 5,200 abstracts were reviewed and it is believed that our multi-faceted search methodology has 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 et al., 1994; Manning, 2004; McCain & Mustard, 1999; Mrazek & Brown, 2002; Nelson et al., 2001; The Developmental Crime Prevention Consortium, 1999; Yoshikawa, 1994), as well as some not previously covered by these reviews.

### **5.3.5: Unit of analysis for relevant intervention programs**

“Early childhood intervention programs” was the unit of analysis for this thesis. This incorporated intervention programs that focused on developing or enhancing child, parent-child, or family wellbeing of at-risk children and their families, and that began before the child commenced primary school. All intervention programs consisted of a treatment group (preschool program participants) and a control group (no preschool participation). In a few cases, both groups received some form of preschool program. When this occurred, only those studies where the intervention program participants received an additional component of a preschool program (e.g. parenting program) were included in the analysis. Studies which

compared interventions of varying scope and intensity were included in the results and discussion.

A total of 11 interventions programs (incorporating 17 follow-up studies) were identified which met the strict criteria discussed above. Table 5.1 provides an overview of these 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. Column 1 of the table identifies the intervention with relevant citation/s. Column 2 provides the research methodology (e.g. randomised design, matched pairs etc.) applied in each study. Column 3 highlights the length of the program in years, and columns 4 and 5 provide the sample sizes for the experimental group and comparison group respectively. Column 6 identifies the child's age at follow-up, and column 7 provides an overview of outcomes that were measured at follow-up.

**Table 5.1: Early intervention programs included in the meta-analysis**

<b>Intervention Citation/Country of study</b>	<b>Study Design (randomised design – yes/no)</b>	<b>Program length</b>	<b>Sample Size Intervention (En)</b>	<b>Sample Size Control (Cn)</b>	<b>Child's age at follow-up</b>	<b>Outcomes</b>
Abecedarian Project (Campbell & Ramey, 1994, 1995b; 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, 2006b; Johnson &	Matched control	5 Years	En =84	Cn =160	Age 13-16	Academic achievement/school performance, mother and family development/family

<b>Intervention Citation/Country of study</b>	<b>Study Design (randomised design – yes/no)</b>	<b>Program length</b>	<b>Sample Size Intervention (En)</b>	<b>Sample Size Control (Cn)</b>	<b>Child's age at follow-up</b>	<b>Outcomes</b>
Blumenthal, 2004b) United States of America						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	Intervention vs.	3 years	*En (group	*Cn (group	Age 15 years	Social/emotional



<b>Intervention Citation/Country of study</b>	<b>Study Design (randomised design – yes/no)</b>	<b>Program length</b>	<b>Sample Size Intervention (En)</b>	<b>Sample Size Control (Cn)</b>	<b>Child's age at follow-up</b>	<b>Outcomes</b>
Infancy Project (Eckenrode et al., 2000b; Olds et al., 1998) United States of America	Control (yes)		3)= 100 *En (group 4)= 116 **En (group 4)= 38	1 & 2)= 184 **Cn = 62		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 et al., 1998) United States of America	Home-based intervention with mothers vs. Control (yes)	>1 year-2 years	Full 2 year En = 70	Less than 2 year Cn = 38 No program n = 15	Age 13, 17 and 22 years	Social-emotional outcomes, cognitive, high school graduation, school drop out
Perry Preschool Program (Berrueta-Clement et al.,	Perry Preschool vs. Control (yes)	2 years	Preschool En = 68	Preschool Cn = 65	Preschool, age 13, 18	Social-emotional, personal behaviour, social development,

<b>Intervention Citation/Country of study</b>	<b>Study Design (randomised design – yes/no)</b>	<b>Program length</b>	<b>Sample Size Intervention (En)</b>	<b>Sample Size Control (Cn)</b>	<b>Child's age at follow-up</b>	<b>Outcomes</b>
1984) United States of America			Age 13 En = 68 Age 18 En = 55	Age 13 Cn = 65 Age 18 Cn = 62		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,

<b>Intervention Citation/Country of study</b>	<b>Study Design (randomised design – yes/no)</b>	<b>Program length</b>	<b>Sample Size Intervention (En)</b>	<b>Sample Size Control (Cn)</b>	<b>Child's age at follow-up</b>	<b>Outcomes</b>
						annual income, self-supporting), effects on health, family, and children in mid-life.
The Syracuse Family Research Development Program (FDRP) (Lally et al., 1988) United States of America	Multi-component vs. comparison (no)	5 years	Preschool En = 82	Preschool Cn = 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	En = 65	Cn = 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)

### 5.3.6: Preschool programs analysed

The eleven intervention programs listed in Table 5.1 constitute a variety of forms of early childhood intervention. These include structured preschool programs (SPP), centre-based childcare/developmental day care programs (CBCC), home visitation programs (HV), family support services (FSS), and parental education (PE). Programs are disaggregated by program components in Chapter 6, Table 6.5. In this section we provide a description of the various forms of intervention.

#### (a) *Structured preschool program (SPP)*

A structured preschool program (SPP), typically centre-based, provides structured types of learning experiences within the context of developmentally orientated programs in which children work together with the preschool teacher to develop their language and cognitive skills. A structured preschool curriculum tends to incorporate mathematics concepts, language arts, reading readiness, social studies and science into structured classroom activities (Coalition for evidence-based policy, 2005; Parks, 2000; Weikart & Schweinhart, 1992). Programs in this meta-analysis which incorporate a SPP component include the Abecedarian project, Chicago Child-Parent Centres, The Early Training Project, Learning to Learn Project, Perry Preschool Project, Parent Child Development Centres, and Project Follow Through.

#### (b) *Centre-based/developmental day care (CBCC)*

Centre-based/developmental day care (CBCC) was developed to meet the growing demands for day care (due to the increasing numbers mothers returning to the

workforce) coupled with the need to provide a developmental preschool program for children that would ultimately increase school readiness and assist children in attaining the social skills to function within society. The aim of developmental day care is to 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). The goal of the majority of developmental day care centres is to:

- provide children with elementary academic readiness to promote optimal cognitive development and the social skills to successfully negotiate the early primary school experience;
- provide the opportunity for children to develop gross motor skills;
- instil a sense of values which will allow children to become responsible and constructive citizens;
- provide parents with appropriate parenting skills and techniques;
- act as a medium between families and various social service agencies; and,
- provide children with numerous cultural and real-life experiences (O'Brien Caughy, DiPietro, & Strobino, 1994; Reynolds, 1994; Reynolds et al., 2001).

Programs in this meta-analysis which incorporate a CBCC component include the Abecedarian Project, Chicago Child-Parent Centres, The Louisville Experiment, and the Syracuse University Family Development Research Program.

### (c) *Home visitation*

Home visitation programs were developed to provide a variety of services to parents to reduce incidences of child abuse and neglect resulting from parents who

lack the skills to properly care for an infant or without the social support networks required to support them as new parents (e.g. relatives or neighbours). Home visitation programs involve either a nurse, social worker, or neighbour visiting the homes of expectant mothers/parents to help motivate them as parents to learn and accept the help of others so that they might provide a loving and nurturing environment for their new child (Olds, Henderson, & Kitzman, 1994). Benefits of the program include:

- linking parents and infants to preventive medical care;
- providing resources to parents who are eager to learn the skills necessary to care for their child;
- promoting parent-child attachment and bonding;
- aiding parents in developing appropriate expectations for their child's development and provide resources to foster that development;
- providing support to families who have other younger children
- identifying and providing guidance and support to already overburdened families; and,
- facilitating the formation of long-term and trusting relationships between families and their support networks (James, 1995; Olds et al., 1997).

Programs in this meta-analysis which incorporate a HV component include the Early Training Project, Elmira Nurse Home Visitation Program, the Perry Preschool Program, the Syracuse University Family Development Research Program, Parent Child Development Centres, and Project Follow Through.

*(d) Family support services*

Family support services (FSS) work in partnership with parents to aid in crisis situations to minimise 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. Family support services tend to offer individual and family counselling, financial management, anger management, domestic abuse, sexual abuse, learning disability, and youth support services (Children's Home Society and Family Services, 2007).

Programs in this meta-analysis which incorporate a FSS component include the Chicago Child-Parent Centres, Elmira Nurse Home Visitation Program, the Mother-Child Home Program, the Syracuse University Family Development Research Program, and Parent Child Development Centres.

*(e) Parental education*

The primary aim of a parent education program is to improve core parenting skills. Sanders (2003) argues that aspects of core parenting skills include:

- improving parents' observation skills – e.g. monitoring both children's and ones own behaviour;
- developing parent-child relationship enhancement skills – e.g. spending quality time, showing affection, and talking with your child;

- encouraging desirable behaviour – e.g. providing activities which are engaging, developing the skill of giving non-verbal engagement, and providing descriptive praise;
- teaching new skills and behaviours – e.g. setting a good example and setting new developmentally appropriate goals;
- managing misbehaviour – e.g. establishing rules, adopting directed discussion; giving clear, concise and calm instructions;
- preventing problems in high-risk situations – e.g. planning and preparation skills, discussing ground rules with the child, providing incentives and discussing consequences of inappropriate behaviour;
- self-regulation skills – e.g. self-evaluation on one's strengths and weaknesses, and setting personal goals;
- mood management and coping skills – e.g. relaxation and stress management skills, developing coping plans for difficult situations, and challenging unhelpful thoughts; and,
- partner support and communication skills – e.g. improving personal communication, having casual conversations, supporting one another during difficult situations, problem solving, and improving relationships.

The program in this meta-analysis which incorporates a PE component is the Parent-Child Development Centre.

### **5.3.7: Variables coded in meta-analysis**

Fourteen variables were coded for each study. A dissemination of program variables is provided in Chapter 6, Tables 6.2, 6.3, and 6.4. Variables include 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. randomised design, no. of outcome measures).

Length of a program was coded in weeks (where there are 9 months of school per year and 4.3 weeks per month), and intensity was coded according to the intended amount of sessions for both parents and children. This approach was taken, given that most studies did not report actual attendances at each session. Therefore, intensity equalled 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. 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.

The coding of studies began by developing a coding system and assigning definitions for each code. Next, variables for all studies were coded. 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.

### **5.3.8: Operationalisation of outcome domains**

A central weakness in previous meta-analyses has been the inconsistency in the operationalisation of outcome measures across primary studies. Recent meta-analytic reviews (Nelson et al., 2001; Nelson et al., 2003) have not clearly justified the methods for their outcome domain selection. However, in this thesis, we have

created domains based on a review of the psychometric literature (e.g. Achenbach, 1991; Achenbach & Rescorla, 2000b; Arnold, O'Leary, Wolff, & Acker, 1993; Caldwell & Bradley, 2001; Coopersmith, 1967, 1975; Corle, Sharbaugh, Mateski, Coyne, Paskett, Cahill, Daston, Lanza, & Schatzkin, 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, 1995b; Campbell et al., 2002). This has been done in order to include dependent variables that may have been defined and/or measured differently across studies and to provide a more accurate and reliable picture of the outcomes associated with all seven domains studied. Table 5.2 provides a description of the seven outcome domains and their relevant indicators identified by the psychometric literature and longitudinal research, and subsequently used in the current study. A detailed description of the relevant indicators is provided in Appendix A.

Initially, more fine-grained outcome domains were considered for adoption. However, due to limited data across interventions it was decided that a smaller set of outcome domains be used and some outcomes be consolidated, resulting in a more meaningful analysis.

**Table 5.2: Outcome domains and operationalisations across studies included in the meta-analysis**

<b>Outcome domains</b>	<b>Operationalisations</b>
<i>Educational success</i>	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.
<i>Cognitive development</i>	IQ; achievement tests; school grades; rating of academic skill and performance; school failure.
<i>Social-emotional development</i>	Parent/teacher rating of problem behaviour; social skills; self-esteem; self-confidence; obsessive-compulsive behaviour.
<i>Deviancy</i>	Rates of delinquent behaviour; drug use (e.g. marijuana and alcohol); lying about age (e.g. false Id); running away from home; caught breaking the law; gang involvement.
<i>Social participation</i>	Casual employment in teen years; socio-economic success; engaged in skilled jobs (e.g. electrician); makes active response to problems.
<i>Criminal justice outcomes</i>	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 behaviour (PINS).
<i>Family wellbeing</i>	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.

It is believed that the seven broad outcome domains, while not perfect, are conceptually distinct and meaningful but also tap into useful outcome domains never before meta-analysed.

The development of outcome domains allowed for an examination of a much broader range of variables than would have otherwise been possible. While efforts were made to keep the constructs homogeneous, this was not always possible. However, care was taken to ensure that dependent variables placed under each domain remained consistent with the psychometric literature (Nelson et al., 2003).

### **5.3.9: Effect size protocols used in this meta-analysis**

Effect sizes that were reported in book chapters, peer reviewed journal articles and major reports were utilised. Effect sizes were pooled, within each study, according to the seven outcome domains in Table 5.2 (educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes and family wellbeing) around both transition periods (transitions from primary to high school and high school to the work force or post secondary education). Moreover, the formulas outlined in Section 5.1 were employed for measuring  $d$ , variance of  $d$ , weighted mean effect size ( $\underline{d}$ ), calculation of 95% confidence intervals, calculation of median effect size, and measuring the  $Q$  statistic.

In the context of the current study, the standardised mean difference effect size for each outcome construct represented the strength (size) and direction (positive or negative) of the difference between treatment and control groups. Various methods were applied to calculate Cohen's  $d$ . This, of course, was dependent upon how

research results were reported (e.g. means and standard deviations, t-tests, F-tests, frequencies, percentages) as discussed in Section 5.1.

Calculation of effect sizes was performed in the context of the current study so that the direction of the outcome favoured treatment groups. Therefore, a positive effect size indicated a deviation between treatment groups and control groups in favour of the treatment groups. A negative effect size, on the other hand, indicated a lower standardised mean in favour of the treatment groups.

Ranges of effect size per outcome construct surrounding a particular transition point are also provided. Finally, studies that reported particular outcomes as “not significant” were translated into an effect size of 0. Previous meta-analyses that have faced this problem, have argued that this approach is a conservative way of translating “non-significant” finding into effect sizes (Durlak & Lipsey, 1991; Durlak & Wells, 1997; Rosenthal, 1995).

### **5.3.10: Research questions**

This meta-analysis, as stated above, aims to identify the salient outcome domains and their indicators during the adolescent life phase. Additionally, it aims to estimate the effectiveness of early childhood intervention programs in terms of effect size on the seven outcome domains identified in Chapter 4, namely educational success, cognitive development, social-emotional development, deviance, social participation, involvement in the criminal justice system, and family wellbeing. The meta-analysis addresses four key questions critical to this thesis, namely:

1. How effective are early childhood intervention programs on the outcome domains (educational success, cognitive development, social-emotional development, deviance, social participation, involvement in criminal justice and family wellbeing) during the adolescent life phase?
2. Do programs that contain a preschool education component (e.g. structured preschool program) have larger effects on the domains of educational success and cognitive development than programs that do not contain a preschool educational component (e.g. family support services and home visitation)?
3. Are effect sizes larger for those programs with a follow through component (e.g. post intervention or supplementary programs to further support the family and child) than for those without?
4. Are the program characteristics, length, intensity, and number of program components important moderators of program success on outcomes during adolescence?

#### **5.4: Conclusion**

In this chapter, we have provided a detailed description of the meta-analytic methodology used in this thesis. Applying the meta-analytic method we will identify best estimates of the effectiveness (in terms of effect size) of early childhood intervention programs on outcomes during adolescence. Moreover, we investigate (together with an analysis of the psychometric test literature) the salient outcome domains (non health-related) considered significant during adolescence. Results of the meta-analytic review will then be used to estimate the effectiveness of early childhood intervention programs in terms of the effect size on a range of outcome domains.

Information on outcome domains, indicators and effect size provides the basis for developing a hierarchical decision-making structure we apply in Chapters 8 and 9. The analytic hierarchy process will be used to assist in selecting between alternative early childhood intervention programs and identify preference scores for non health-related outcomes during adolescence.

Chapter 6 discusses the results of our meta-analysis identifying the effects of 5 forms of preschool intervention (structured preschool program (SPP), home visitation (HV), centre-based childcare/developmental day care (CBCC), family support services (FSS) and parental education (PE)) on at-risk children's educational success (ES), cognitive development (CD), social-emotion development (SED), deviancy (D), social participation (SP), criminal justice outcomes (CJ), and family wellbeing (FW) during adolescence (age 12-19 years).

## **CHAPTER 6: RESULTS: A META-ANALYSIS OF LONGITUDINAL RESEARCH ON THE ADOLESCENT OUTCOMES ASSOCIATED WITH EARLY CHILDHOOD INTERVENTION PROGRAMS**

### **6.1: Introduction**

This chapter provides results of the meta-analysis and discusses the effects of 5 forms of preschool intervention (structured preschool program (SPP), home visitation (HV), centre-based childcare/developmental day care (CBCC), family support services (FSS) and parental education (PE)) on at-risk children's educational success (ES), cognitive development (CD), social-emotional development (SED), deviancy (D), social participation (SP), criminal justice outcomes (CJ), and family wellbeing (FW) during adolescence (age 12-19 years). This study does not review secondary prevention or indicated prevention programs for children who are already demonstrating early signs of problem behaviour. In Section 6.2, descriptive statistics are provided. Next, ranges and weighted mean  $d$ s, corrected for small sample size, for the seven outcome domains are presented. This is followed by a discussion of the ranges of effect size for the seven outcome domains. Next, three program characteristics, length, intensity, and number of program components are analysed and discussed with respect to effects on outcomes during adolescence. A discussion of the meta-analysis results follows. We then discuss methods of moving forward to improve our understanding and knowledge of the effects of early childhood intervention projects within and beyond adolescence. Finally, we discuss methods of improving how policy-makers make decisions under the veil of risk and uncertainty and, how results of our meta-analysis can be applied to develop a methodology to



improve on current economic analyses techniques applied in the area of criminal justice research.

## **6.2: Meta-analytic results - coding and descriptive statistics**

A total of 17 studies were found that examined the effects of early childhood intervention programs on at-risk children and their families during the adolescent life phase. Table 6.1 outlines the codes (column 1) assigned to those studies (column 2) included in the meta-analysis. The sample of studies included in this analysis represents the follow-up analyses of the eleven independent intervention projects described in Chapter 5 (Section 5.3.5, Table 5.1). Column 1 of Table 5.1 highlights the correspondence between the included studies and the associated projects.

Table 6.2 provides descriptive information on program characteristics. Early childhood intervention programs that contained a structured preschool program (SPP) made up over sixty percent (63.6%) of the total studies included in the meta-analysis. Over forty-five percent of studies included in the analysis incorporated a home visitation (HV) component, and a family support (FSS) component. Further, thirty six percent of programs included a centre-based childcare/developmental day care component. Finally, only nine percent of total programs included in the meta-analysis incorporated a parental education (PE) component. Seventy three percent (8 out of eleven) of programs included in the meta-analysis incorporated 1-2 intervention components (e.g. SPP or CBCC). Only three early intervention programs included in the meta-analysis incorporated three or more program components into the intervention (e.g. SPP, HV and PE). Thirty six percent of programs included into the meta-analysis had duration (length of program) of longer than three years. Over forty-

five percent of programs had an intensity of five hundred sessions; and finally, over thirty-six percent of programs contained a follow-through intervention.

**Table 6.1: Codes for individual (primary) studies included in the meta-analysis.**

Study Code	Project	Study
1	Abecedarian Project	Campbell, Ramey, Pungello, Sparling, & Miller-Johnson (2002)
2	Abecedarian Project	Campbell & Ramey (1994)
3	Abecedarian Project	Campbell & Ramey (1995a)
4	Chicago Child-Parent Centre	Reynolds, Temple, Robertson, & Mann (2001)
5	Chicago Child-Parent Centre	Reynolds (1994)
6	Early Training Project	Lazar, Darlington, Murray, Royce, Snipper, & Ramey (1982)
7	Early Training Project	Gray & Klaus (1970)
8	Elmira Prenatal/ Early Infancy Project	Eckenrode, Ganzel, Henderson, Smith, Olds, Powers, Cole, Kitzman, & Sidora (2000a)
9	Learning to Learn Project	Sprigle & Shaefer, L. (1985)
10	Louisville Experiment	Miller & Bizzell (1983)
11	Mother-Child Home Program	Levenstein, Levenstein, Shiminski, & Stolzberg (1998)
12	Perry Preschool Program	Berrueta-Clement, Schweinhart, Barnett, Epstein & Weikart (1984)
13	The Syracuse Family Research Development Program	Lally, Mangione, & Honig (1988)
14	Parent-Child Development Centres	Johnson (2006a)
15	Parent-Child Development Centres	Johnson, & Blumenthal, (2004a)
16	Direct Instruction project	Meyer (1984)
17	Elmira Prenatal/ Early Infancy Project	Olds, Henderson, Cole, Eckenrode, Kitzman, Luckey, Pettitt, Sidora, Morris, & Powers (1998)

**Table 6.2: Program characteristics of included studies**

<b>Program Characteristics (n=11)</b>	<b>Number</b>	<b>% of total programs</b>
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 centre-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

\*Intensity (number of sessions) = contacts per week x total number of program weeks

Tables 6.3 and 6.4 provide further descriptive information regarding the characteristics of program participants and study characteristics. Results highlight that over seventy percent of participants who made up both control and experimental groups in the early intervention programs included in the meta-analysis were predominantly from African-American ethno-racial backgrounds. Moreover, approximately sixty-five percent of studies (n=17) included in the meta-analysis

utilised a randomised design, with over eighty percent of sample sizes greater than 300.

**Table 6.3: Characteristics of program participants**

<b>Characteristics of program participants (n=11)</b>		
• Predominately African-American ethno-racial background ( $\geq 50\%$ participants African-American)	8	72.3
• Other ethno-racial backgrounds (<50% participants African-American e.g. Mexican-American)	3	27.3

**Table 6.4: Study Characteristics**

<b>Study Characteristics (n=17)</b>		
• Randomised design	11	64.7
• Non-randomised (e.g. matched-pairs)	6	35.3
Target population clearly described	17	100
• Sample size		
• <300	14	82.4
• $300 \leq 500$	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 6.5 presents weighted mean  $d$ s (effect size), corrected for small sample size. The weighted mean effect size was computed by weighting individual effect sizes by the inverse of its variance. A general formula for calculating a weighted mean effect size is demonstrated in Lipsey and Wilson (Lipsey & Wilson, 2000, pp. 113-114). The variance and inverse variance, and confidence intervals lower and upper (95%) corrected for small sample size are presented for the seven outcome domains (Educational success (ES), cognitive development (CD), social-emotional development (SED), deviance (D), social participation (SP), criminal justice (CJ), and family wellbeing (FW)) during the adolescent life phase. Column 1 in Table 6.5

identifies the intervention program and the age at which the follow-up occurred. Column 2 identifies the coding for the study number; a total of seventeen studies are included into the meta-analysis. Column 3 identifies the various program components that were used in the intervention (e.g. structured preschool component (SPP), home visitation component (HV)). Further, Columns 4-10 incorporate the results of the intervention projects on the above-mentioned outcome domains. Appendix B provides individual effect sizes, variances and inverse variances for early childhood intervention programs included in the meta-analysis including dependent variables (e.g. reduction in special education, reduction in grade retention) measured under each outcome domain (e.g. educational success (ES), cognitive development (CD)).

**Table 6.5: Effect sizes (d and d.), variance, inverse variance, CI Lower and Upper (95%), corrected for small sample size, for the seven outcome domains during the adolescent life phase.**

			Average effect sizes ( <u>d</u> ), Sample size ( <i>En</i> , <i>Cn</i> ), variance of ( <u>d</u> ), inverse variance, CI Lower (95%), CI Upper (95%)						
Program + (age at follow-up)	Study No.	Component (Major component)*	Educational success (ES)	Cognitive development (CD)	Social- emotional development (SED)	Deviancy (D)	Social participation (SP)	Criminal justice outcomes (CJ)	Family well- being (FW)
Abecedarian Project (20 Yrs)	1	CBCC (SPP) Preschool (5 years) vs. no preschool control	<u>d</u> = 0.44 <i>En</i> = 53, <i>Cn</i> = 51 Var ( <u>d</u> ) = 0.039 Inv Var = 25.380 CI Low = 0.0496 CI Up= 0.828	<u>d</u> = 0.47 <i>En</i> = 53, <i>Cn</i> = 51 Var ( <u>d</u> ) = 0.039 Inv Var = 25.296 CI Low = 0.079 CI Up= 0.858		<u>d</u> = 0.49 <i>En</i> = 53, <i>Cn</i> = 51 Var ( <u>d</u> ) = 0.039 Inv Var = 25.241 CI Low = 0.097 CI Up= 0.878	<u>d</u> = 0.43 <i>En</i> = 53, <i>Cn</i> = 51 Var ( <u>d</u> ) = 0.039 Inv Var = 25.415 CI Low = 0.036 CI Up= 0.814	<u>d</u> = 0.21 <i>En</i> = 53, <i>Cn</i> = 51 Var ( <u>d</u> ) = 0.039 Inv Var = 25.843 CI Low = -0.172 CI Up= 0.599	
Abecedarian Project (12 Yrs)	2	CBCC (SPP) 8 years of intervention – (5 yrs from infancy to kindergarten and 3 years in primary grades) vs. control	<u>d</u> = 0.79 <i>En</i> = 25, <i>Cn</i> = 23 Var ( <u>d</u> ) = 0.0197 Inv Var = 10.895 CI Low = 0.196 CI Up= 1.834	<u>d</u> = 0.40 <i>En</i> = 25, <i>Cn</i> = 22 Var ( <u>d</u> ) = 0.0872 Inv Var = 11.470 CI Low = -0.982 CI Up= 0.176					
Abecedarian Project (12 Yrs)	2	CBCC (SPP) Preschool (5 years) vs. no preschool control	<u>d</u> = 0.24 <i>En</i> = 24, <i>Cn</i> = 23 Var ( <u>d</u> ) = 0.08576 Inv Var = 11.661 CI Low = -0.334 CI Up= 0.814	<u>d</u> = 0.53 <i>En</i> = 22, <i>Cn</i> = 22 Var ( <u>d</u> ) = 0.094 Inv Var = 10.634 CI Low = -1.126 CI Up= 0.076					
Abecedarian Project (15 Yrs)	3	CBCC (SPP) 8 years of intervention – (5 yrs from infancy to kindergarten and 3 years in primary grades) vs. control	<u>d</u> = 0.69 <i>En</i> = 25, <i>Cn</i> = 23 Var ( <u>d</u> ) = 0.088 Inv Var = 11.308 CI Low = 0.107 CI Up= 1.272	<u>d</u> = 0.59 <i>En</i> = 25, <i>Cn</i> = 23 Var ( <u>d</u> ) = 0.087 Inv Var = 11.474 CI Low = -1.173 CI Up= -0.015					

Abecedarian Project (15 Yrs)	3	CBCC; (SPP) Preschool (5 years) vs. no preschool control	$\underline{d} = 0.74$ $En = 23, Cn = 23$ $Var(\underline{d}) = 0.093$ $Inv Var = 10.757$ $CI Low = -1.341$ $CI Up = -0.146$					
Chicago Child-Parent Center (CPC) (20 Yrs)	4	(SPP); CBCC; FSS Preschool (6 yrs – preschool to grade 3) vs. control	$\underline{d} = 0.16$ $En = 837, Cn = 444$ $Var(\underline{d}) = 0.003$ $Inv Var = 289.269$ $CI Low = 0.045$ $CI Up = 0.275$					$\underline{d} = 0.20$ $En = 837, Cn = 444$ $Var(\underline{d}) = 0.003$ $Inv Var = 288.799$ $CI Low = 0.085$ $CI Up = 0.315$
Chicago Child-Parent Center (CPC) (11 Yrs)	5	(SPP); CBCC; FSS Preschool only vs. control	$\underline{d} = -0.12$ $En = 207, Cn = 191$ $Var(\underline{d}) = 0.010$ $Inv Var = 99.161$ $CI Low = -0.317$ $CI Up = 0.077$	$\underline{d} = 0.18$ $En = 207, Cn = 191$ $Var(\underline{d}) = 0.010$ $Inv Var = 98.939$ $CI Low = -0.017$ $CI Up = 0.377$	$\underline{d} = 0.12$ $En = 207, Cn = 191$ $Var(\underline{d}) = 0.003$ $Inv Var = 289.645$ $CI Low = 0.005$ $CI Up = 0.2351$			$\underline{d} = 0.11$ $En = 207, Cn = 191$ $Var(\underline{d}) = 0.010$ $Inv Var = 99.189$ $CI Low = -0.087$ $CI Up = 0.307$
Chicago Child-Parent Center (CPC) (11 Yrs)	5	(SPP); CBCC; FSS Preschool (6 yrs – preschool to grade 3) vs. control	$\underline{d} = -0.39$ $En = 160, Cn = 191$ $Var(\underline{d}) = 0.012$ $Inv Var = 85.454$ $CI Low = -0.602$ $CI Up = -0.178$	$\underline{d} = 0.29$ $En = 160, Cn = 191$ $Var(\underline{d}) = 0.011605$ $Inv Var = 86.16676$ $CI Low = 0.078852$ $CI Up = 0.501148$	$\underline{d} = 0.11$ $En = 207, Cn = 191$ $Var(\underline{d}) = 0.003$ $Inv Var = 289.711$ $CI Low = -0.005$ $CI Up = 0.225$			$\underline{d} = 0.11$ $En = 160, Cn = 191$ $Var(\underline{d}) = 0.012$ $Inv Var = 86.935$ $CI Low = -0.100$ $CI Up = 0.320$
Early Training Project (18 Yrs)	6	(SPP); HV Preschool vs. control	$\underline{d} = 0.29$ $En = 36, Cn = 19$ $Var(\underline{d}) = 0.081$ $Inv Var = 12.321$ $CI Low = -0.270$ $CI Up = 0.847$	$\underline{d} = 0.16$ $En = 34, Cn = 18$ $Var(\underline{d}) = 0.085$ $Inv Var = 11.735$ $CI Low = -0.412$ $CI Up = 0.732$	$\underline{d} = 0.65$ $En = 24, Cn = 26$ $Var(\underline{d}) = 0.054$ $Inv Var = 18.444$ $CI Low = -1.110$ $CI Up = -0.197$			
Early Training Project (13 Yrs)	6	(SPP); HV Preschool vs. control		$\underline{d} = 0.46$ $En = 41, Cn = 18$ $Var(\underline{d}) = 0.082$ $Inv Var = 12.239$ $CI Low = -0.105$ $CI Up = 1.015$				

Early Training Project (11 Yrs)	7	(SPP); HV Preschool vs. control		$\underline{d} = 0.51$ $En = 38, Cn = 41$ $Var(\underline{d}) = 0.052$ $Inv Var = 19.098$ $CI Low = 0.063$ $CI Up = 0.960$					
Elmira Nurse Home Visitation Program (15 Yrs)	8	(HV); FSS Program vs. control							$\underline{d} = 0.38$ $En = 116, Cn = 184$ $Var(\underline{d}) = 0.014$ $Inv Var = 69.935$ $CI Low = 0.148$ $CI Up = 0.617$
Learning to Learn (12 Yrs)	9	(SPP) Learning to Learn (preschool through grade 1) vs. Head Start	$\underline{d} = 0.94$ $En = 21, Cn = 19$ $Var(\underline{d}) = 0.111$ $Inv Var = 8.977$ $CI Low = 0.290$ $CI Up = 1.598$	$\underline{d} = 0.51$ $En = 21, Cn = 19$ $Var(\underline{d}) = 0.103$ $Inv Var = 9.662$ $CI Low = -1.140$ $CI Up = 0.121$					
Louisville experiment (13 Yrs)	10	(CBCC) Bereiter-Engelmann Preschool vs. Control		$\underline{d} = 0.24$ $En = 13, Cn = 14$ $Var(\underline{d}) = 0.149$ $Inv Var = 6.691$ $CI Low = -1.001$ $CI Up = 0.514$					
Mother-Child Home program (17 Yrs)	11	(FSS) Program vs. control	$\underline{d} = 0.34$ $En = 104, Cn = 13$ $Var(\underline{d}) = 0.087$ $Inv Var = 11.491$ $CI Low = -0.241$ $CI Up = 0.915$						
Perry Preschool Program (19 Yrs)	12	(SPP); HV Program vs. control				$\underline{d} = 0.645$ $En = 58, Cn = 63$ $Var(\underline{d}) = 0.03486$ $Inv Var = 28.686$ $CI Low = 0.284$ $CI Up = 1.016$		$\underline{d} = 0.41$ $En = 58, Cn = 63$ $Var(\underline{d}) = 0.034$ $Inv Var = 29.578$ $CI Low = 0.049$ $CI Up = 0.770$	
The Syracuse University Family Development Research Program (15 Yrs)	13	HV; CBCC; (FSS) Program vs. control	$\underline{d} = 0.82$ $En = 82, Cn = 74$ $Var(\underline{d}) = 0.028$ $Inv Var = 35.889$ $CI Low = 0.493$ $CI Up = 1.147$		$\underline{d} = 0.40$ $En = 45, Cn = 39$ $Var(\underline{d}) = 0.049$ $Inv Var = 20.492$ $CI Low = -0.036$ $CI Up = 0.829699$		$\underline{d} = 0.45$ $En = 45, Cn = 39$ $Var(\underline{d}) = 0.049$ $Inv Var = 20.374$ $CI Low = 0.018$ $CI Up = 0.887$	$\underline{d} = 0.48$ $En = 65, Cn = 54$ $Var(\underline{d}) = 0.035$ $Inv Var = 28.677$ $CI Low = 0.114$ $CI Up = 0.846$	$\underline{d} = 0.43$ $En = 51, Cn = 42$ $Var(\underline{d}) = 0.044$ $Inv Var = 22.517$ $CI Low = 0.017$ $CI Up = 0.843$



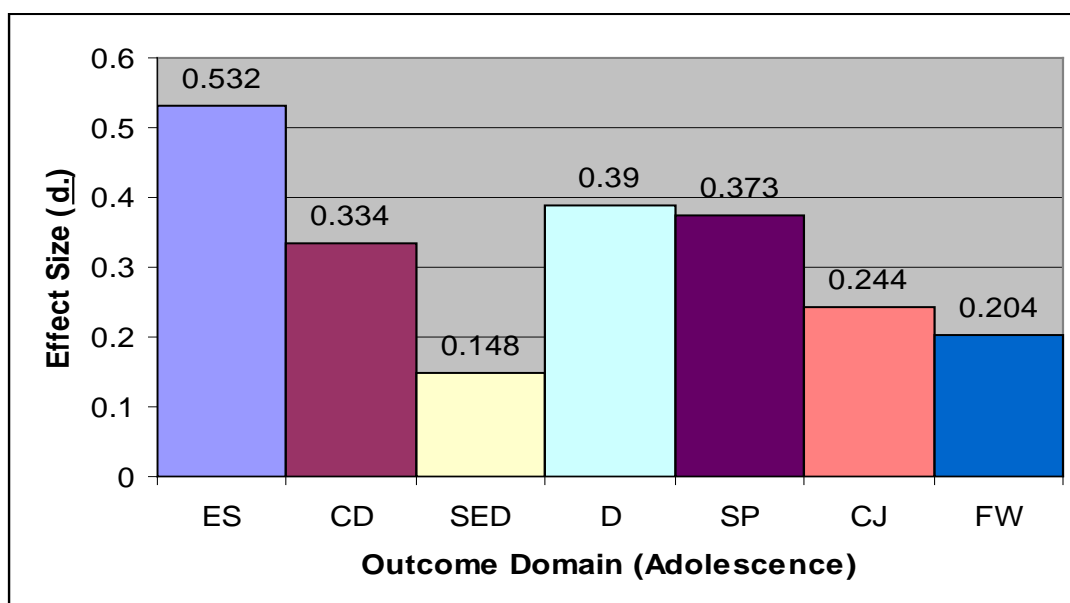
Parent-Child Development Centres (PCDCs) (15 Yrs)	14	(SPP); HV; FSS; PE Houston PCDC vs. control 5 yr program infancy – 5 yrs			$\underline{d} = 0.23$ $En = 84, Cn = 160$ $Var(\underline{d}) = 0.018$ $Inv Var = 54.766$ $CI Low = -0.039$ $CI Up = 0.491$				
Parent-Child Development Centres (PCDCs) (12 Yrs)	15	(SPP); HV; FSS; PE Houston PCDC vs. control 5 yr program infancy – 5 yrs	$\underline{d} = 0$ $En = 84, Cn = 160$ $Var(\underline{d}) = 0$ $Inv Var =$ $CI Low =$ $CI Up =$	$\underline{d} = 0.36$ $En = 89, Cn = 155$ $Var(\underline{d}) = 0.018$ $Inv Var = 55.704$ $CI Low = 0.097$ $CI Up = 0.622$	$\underline{d} = 0$ $En = 84, Cn = 160$ $Var(\underline{d}) = 0$ $Inv Var =$ $CI Low =$ $CI Up =$		$\underline{d} = 0.32$ $En = 89, Cn = 155$ $Var(\underline{d}) = 0.018$ $Inv Var = 55.859$ $CI Low = 0.061$ $CI Up = 0.586$		$\underline{d} = 0$ $En = 84, Cn = 160$ $Var(\underline{d}) = 0$ $Inv Var =$ $CI Low =$ $CI Up =$
Project Follow Through (Brooklyn Project)	16	(SPP); HV Direct Instruction vs. control Preschool through grade 3	$\underline{d} = 0.38$ $En = 65, Cn = 100$ $Var(\underline{d}) = 0.026$ $Inv Var = 38.726$ $CI Low = 0.065$ $CI Up = 0.695$	$\underline{d} = 0.42$ $En = 65, Cn = 100$ $Var(\underline{d}) = 0.026$ $Inv Var = 38.572$ $CI Low = -0.738$ $CI Up = -0.107$					
Elmira Nurse Home Visitation Program (15 Yrs)	17	(HV); FSS Program vs. control	$\underline{d} = 0.46$ $En = 38, Cn = 62$ $Var(\underline{d}) = 0.0435$ $Inv Var = 22.987$ $CI Low = 0.0511$ $CI Up = 0.8688$			$\underline{d} = 0.28$ $En = 38, Cn = 62$ $Var(\underline{d}) = 0.0428$ $Inv Var = 23.344$ $CI Low = -0.126$ $CI Up = 0.0686$		$\underline{d} = 0.32$ $En = 38, Cn = 62$ $Var(\underline{d}) = 0.043$ $Inv Var = 23.279$ $CI Low = -0.086$ $CI Up = 0.726$	
<b>Range <math>\underline{d}</math></b>			Range $\underline{d} =$ 0.00-0.94	Range $\underline{d} =$ 0.16-0.74	Range $\underline{d} =$ 0.00-0.65	Range $\underline{d} =$ 0.28-0.65	Range $\underline{d} =$ 0.32-0.45	Range $\underline{d} =$ 0.20-0.48	Range $\underline{d} =$ 0-0.43
<b>Mean <math>\underline{d}</math></b>			$\underline{d} = 0.532$	$\underline{d} = 0.334$	$\underline{d} = 0.148$	$\underline{d} = 0.39$	$\underline{d} = 0.373$	$\underline{d} = 0.244$	$\underline{d} = 0.204$

\*SPP = Structured preschool program; CBCC = Centre-based childcare/ developmental day care; HV = home visitation; FSS = family support services; and PE = parental education.

### 6.3: Mean effect size and Q statistic results

Figure 6.1 summarises the mean effect sizes relating to the effect of early childhood intervention programs on seven outcome domains measured throughout adolescence. The mean weighted effect size ( $d$ ) for outcome domains CJ and FW are 0.244 and 0.204 respectively, highlighting small effects. Mean effect sizes for the outcome domains CD, SP and D were 0.334, 0.373 and 0.39 respectively, which indicates a small to medium effect. The largest effect size is displayed on the domain ES (0.532), which demonstrates a medium to medium-high effect, while the SED domain demonstrates the smallest effect of (0.148).

**Figure 6.1: Weighted average effect sizes ( $d$ ) corrected for sample size for seven adolescent outcomes**



Effect sizes have been adjusted for sample size. Three outliers (under the ES domain) were identified and removed from the analysis. The outliers came from the Chicago Child-Parent Centre 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. However, in this study we chose to remove outliers, as proposed by Lipsey and Wilson (2001).

They state:

The purpose of meta-analysis is to arrive at a reasonable summary of the quantitative findings of a body of research studies. The purpose is usually not well served by the inclusion of extreme effect size values that are noticeably discrepant from the preponderance of those found in the research of interest and, hence, unrepresentative of the results of that research...extreme effect size values have disproportionate influence on the values of the means, variances, and other statistics used in the meta-analysis and may distort them in misleading ways...One common procedure for handling a few outliers that are not believed to be representative of study findings is to simply eliminate them from the effect size distribution (pp. 107-108).

The overall Q test for homogeneity was not significant ( $p < 0.001$  and  $p < 0.05$ ) for all outcome domains. In all cases the null hypothesis of homogeneity was not rejected. Table 6.6 provides results of the Q test highlighting the weighted mean effect size ( $\underline{d}$ ), the variance of ( $\underline{d}$ ), confidence intervals lower and upper (95%), the Q test result and the degrees of freedom. Results demonstrate that there were no outliers among the effect sizes included in the analysis. A detailed description of the Q test is provided in Chapter 5.

**Table 6.6: Results - Test of homogeneity for seven outcome domains**

	<b>ES</b>	<b>CD</b>	<b>SED</b>	<b>D</b>	<b>SP</b>	<b>SR</b>	<b>FO</b>
<b><u>d.</u></b>	0.532	0.334	0.148	0.39	0.373	0.244	0.204
<b>Var <u>d.</u></b>	0.005	0.003	0.001	0.021	0.009	0.003	0.004
<b>CI Lower</b>	0.390	0.233	0.072	0.108	0.179	0.145	0.086
<b>95%</b>							
<b>CI Upper</b>	0.675	0.434	0.223	0.670	0.568	0.342	0.321
<b>95%</b>							
<b>Q</b>	8.846	7.319	6.962	0.537	0.360	3.136	4.960
<b>df</b>	10	13	5	2	2	4	4

Note: ES = Educational success, CD = Cognitive development, SED = Social-emotional development, D = Deviancy, SP = Social participation, CJ = Criminal justice outcomes, FW = Family wellbeing.

### **6.3.1: Discussion - Weighted mean effect size**

Results of this meta-analysis demonstrate that early childhood intervention projects have enduring positive effects on the 7 outcome domains (ES, CD, SED, D, SP, CJ, and FW) during the adolescent life phase. The overall size of the effects ( $\underline{d} = 0.35$ ) is in the small to medium range. Converting this to a percentile, this effect size highlights that outcomes for the intervention group sample exceeds by 63% those in the control group sample (see Lipsey & Wilson, 2000, p.153, Table 8.1).

Results of this analysis further demonstrate that the effects of early childhood intervention programs on the outcome domains educational success ( $\underline{d} = 0.53$ ), deviance ( $\underline{d} = 0.39$ ), social participation ( $\underline{d} = 0.37$ ), and cognitive development ( $\underline{d} = 0.33$ ) are greater than on the domains criminal justice outcomes ( $\underline{d} = 0.24$ ), family wellbeing ( $\underline{d} = 0.20$ ), and social-emotional development ( $\underline{d} = 0.15$ ). This conclusion seems consistent with findings presented by Nelson et al. (2003) who found that the

effect of preschool interventions from kindergarten to the start of adolescence (Grade 8) on the outcome domain cognitive development were approximately 0.30. However, Nelson and his colleagues presented a much higher effect (0.30) on the outcome domain family wellness/outcomes. This can be attributed to the fact that Nelson's results incorporated findings from kindergarten to grade 8; and consequently didn't incorporate the results of follow-ups beyond this period. Nelson and his colleagues also presented a much larger effect (0.33) on the outcome domain social-emotional development during the adolescent period (high school + in their terminology). However, this domain was particularly broad given that it incorporated items such as parent and teacher ratings of children's behaviour, social skills, self reported self-esteem, grade retention, placement in special education classes, teenage employment, educational success (e.g. graduation), and criminal behaviour during adolescence. Conversely, our analysis defined social-emotional development in less broad terms and included items such as parent/teacher rating of problem behaviour; social skills; self-esteem; self-confidence; and, obsessive-compulsive behaviour.

### **6.3.2: Ranges of effect size**

Analysing the ranges of effect sizes for the seven outcomes domains during adolescence demonstrated that effect sizes overall ranged from -0.39 to 0.94 (Min/Max). Table 6.7 provides the range of effect size, minimum and maximum, skewness, unweighted mean, standard deviation and sample variance for each outcome domain during adolescence.

**Table 6.7: Statistics for ranges of effect size during adolescence**

	<b>ES</b>	<b>CD</b>	<b>SED</b>	<b>D</b>	<b>SP</b>	<b>CJ</b>	<b>FW</b>
<b>Range (r)</b>	1.33	0.58	0.65	0.37	0.13	0.28	0.43
<b>Min/Max</b>	-0.39/0.94	0.16/0.74	0.00/0.65	0.28/0.65	0.32/0.45	0.20/0.48	0.00/0.43
<b>Skewness</b>	-0.30281	-0.10912	1.01729	-0.40092	-1.57434	0.23613	0.037623
<b>Unweighted mean (d)</b>	0.36	0.43	0.25	0.47	0.40	0.32	0.21
<b>Standard deviation</b>	0.376	0.160	0.237	0.185	0.07	0.123	0.188
<b>Sample variance</b>	0.1416	0.0256	0.0564	0.0344	0.0049	0.0150	0.0353

The outcome domain deviancy demonstrated the largest range (r) (r=0.37; min 0.28-max 0.65; Unweighted Mean=0.47) followed by cognitive development (r=0.58; min 0.16-max 0.74; Mean=0.43), social participation (r=0.13; min 0.32-max 0.45; Mean=0.40), educational success (r=1.33; min -0.39-max 0.94; Mean=0.36), criminal justice outcomes (r=0.28; min 0.20-max 0.48; Mean=0.32), social-emotional development (r=0.65; min 0.0-max 0.65; Mean=0.25), and family wellbeing (r=0.43; min 0.0-max 0.43; Mean=0.21). When analysing unweighted means and ranges of effect size, we found that overall outcomes with respect to effect sizes during adolescence changed in comparison to the use of weighted means effect sizes. Consequently, results highlight that early childhood intervention projects have the greatest effect on the outcome domain deviancy, followed by cognitive development, social participation, educational success, criminal justice outcomes, social-emotional development, and family wellbeing during adolescence. These findings seem to be consistent, although not directly comparable (due to issues discussed in Chapter 4) with previous meta-analyses (e.g. Farrington & Welsh, 2003), which found that overall family-based crime prevention programs displayed a mean effect size of .321 on delinquency outcomes.

### 6.3.3: Effect sizes for ES and CD during adolescence for programs with and without a preschool education component.

During adolescence, the average weighted  $\underline{d}$  on the ES domain for programs with a preschool component ( $n= 8$ ,  $\underline{d} = 0.48$ ,  $\text{Var } \underline{d} = 0.008$ ; CI Lower 95% = 0.299, CI Upper 95% = 0.658,  $Q= 4.989$ ) was lower than that for programs that did not contain a preschool educational component ( $n= 3$ ,  $\underline{d} = 0.62$ ,  $\text{Var } \underline{d} = 0.014$ ; CI Lower 95% = 0.390, CI Upper 95% = 0.857,  $Q= 2.923$ ). However, the range of (d), or unweighted mean, for programs that did contain an educational component was from 0 – 0.94 (Mean = 0.47; SD = 0.313; Sample Var = 0.098), while the range of (d), or unweighted mean, for programs that did not contain a preschool educational component was 0.34 – 0.82 (Mean = 0.58; SD = 0.339; Sample Var = 0.115). In this case, the weighted  $\underline{d}$  tends to overestimate the effectiveness of programs that did not contain a preschool educational component, and as such, more emphasis should be placed on the ranges of effect sizes. Moreover, one should consider the difference between (n) for programs that do and do not contain a preschool educational component.

The average weighted  $\underline{d}$  for the outcome domain cognitive development with a preschool component ( $k= 13$ ,  $\underline{d} = 0.34$ ,  $\text{Var } \underline{d} = 0.002$ ; CI Lower 95% = 0.246, CI Upper 95% = 0.441,  $Q= 7.690$ ) was higher than that for programs that did not contain a preschool educational component ( $k= 2$ ,  $\underline{d} = 0.23$ ,  $\text{Var } \underline{d} = 0.058$ ; CI Lower 95% = -0.239, CI Upper 95% = 0.707,  $Q= 0.0004$ ). The range of (d) for programs that did contain an educational component was from 0.16 – 0.74 (Mean = 0.43; SD = 0.160; Sample Var = 0.026), while the range of (d) for programs that did not contain a preschool educational component was 0.23 – 0.24 (Mean = 0.235; SD = 0.007).

Our meta-analysis suggests that early childhood intervention programs that incorporate a structured preschool or centre-based educational component provide enduring positive effects on the outcome domains of educational success (0.48) and cognitive development (0.34) throughout the adolescent life phase. Results are considerably higher for educational success than cognitive development. This finding seems consistent with the argument made by Nelson et al (2003) that the effects of preschool educational programs on cognitive development are more pronounced during the preschool and early primary school years. However, Nelson and his colleagues do recognise that although the effects on cognitive development from primary school to Grade 8 are smaller than during the preschool year follow-ups, the effects are nevertheless there. Both our findings and those by Nelson and his colleagues are in contrast to those offered by Cohen and Radford (1999) who claim that there is little evidence that programs directed towards families (in particular, mothers and their children) have impact on children. What sets this analysis apart from other analyses has been our ability to demonstrate that programs that incorporate a structured preschool or centre-based educational component into their curriculum do produce good effect sizes (ranging from medium-small to medium effects) on both educational success and cognitive development well into late adolescence and early adulthood. Therefore, this confirms the continual benefits of such program components beyond the short-term and into the medium- to long-term.

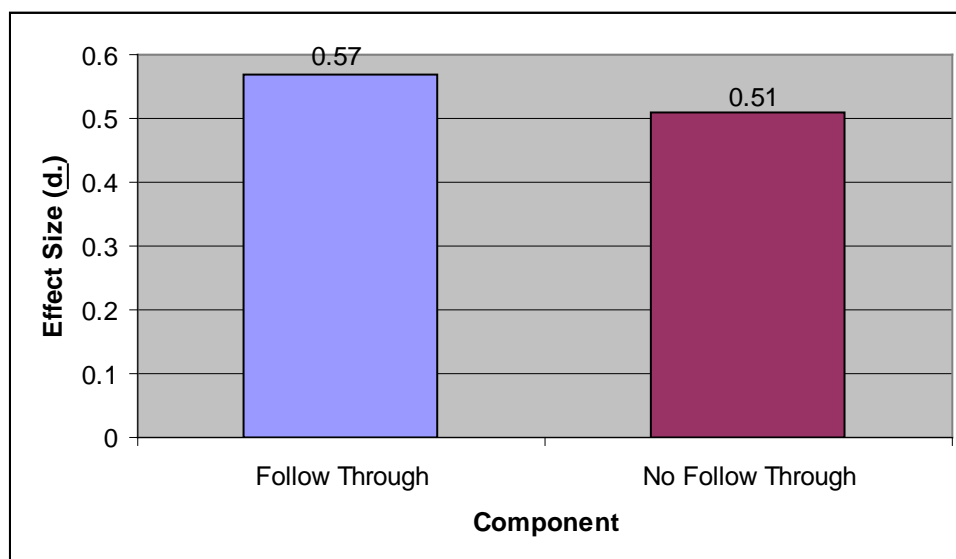
#### **6.3.4: Effect sizes for early childhood intervention programs with and without follow through**

The average weighted  $\underline{d}$  for the ES domain of programs with a follow through component ( $n = 4$ ;  $d = 0.57$ ;  $Var = 0.014$ ;  $CI\ Lower = 0.332$ ;  $CI\ Upper = 0.800$ ;  $Q =$



3.316  $p < .001$ ) was slightly larger than for those programs without a follow through component ( $k = 7$ ;  $d = 0.51$ ;  $Var = 0.008$ ;  $CI\ Lower = 0.334$ ;  $CI\ Upper = 0.692$ ;  $Q = 5.408$ ,  $p < .001$ ) (see Figure 6.2). The range of  $d$ , or unweighted mean, for programs that did contain a follow through component was from 0.38 – 0.94 (Mean = 0.7;  $SD = 0.237$ ; Sample  $Var = 0.056$ ), while the range of  $d$  for programs that did not contain a preschool educational component was 0.0 – 0.82 (Mean = 0.37;  $SD = 0.251$ ; Sample  $Var = 0.063$ ).

**Figure 6.2: Average weighted effect sizes for the outcome educational success (ES) for programs with and without a follow through component**

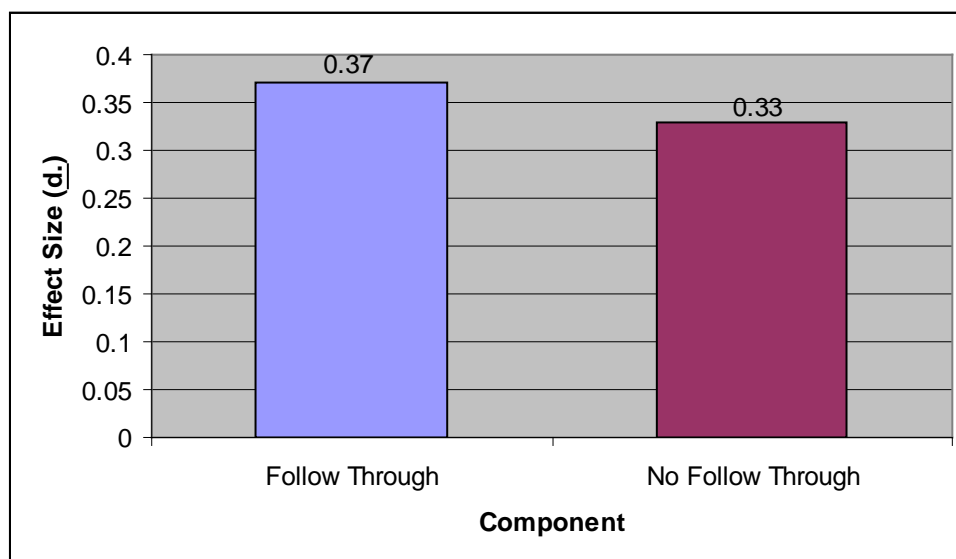


The average weighted  $d$ , for the cognitive development (CD) outcome domain of programs with a follow through component ( $n = 5$ ;  $d = 0.37$ ;  $Var = 0.006$ ;  $CI\ Lower = 0.209$ ;  $CI\ Upper = 0.521$ ;  $Q = 1.399$   $p < .001$ ) was slightly larger than for those programs without a follow through component ( $n = 9$ ;  $d = 0.33$ ;  $Var = 0.004$ ;  $CI\ Lower = 0.203$ ;  $CI\ Upper = 0.451$ ;  $Q = 6.222$ ,  $p < .001$ ) (see Figure 6.3). The range of ( $d$ ), or unweighted mean, for programs that did contain a follow through component was from 0.29 – 0.59 (Mean = 0.442;  $SD = 0.114$ ; Sample  $Var = 0.013$ ),

while the range of (d) for programs that did not contain a preschool educational component was 0.16 – 0.74 (Mean = 0.40; SD = 0.189; Sample Var = 0.036).

Insufficient data was available to measure the effects of follow through versus no follow through on the outcome domains social-emotional development, deviance, social participation, criminal justice and family wellbeing.

**Figure 6.3: Average weighted effect sizes for the outcome cognitive development (CD) for programs with and without a follow through component**



The meta-analysis revealed that programs with a follow through component into the early primary school years (e.g. preschool to Grade 3) had strong effects on the outcomes educational success (.57) and cognitive development (.37). This finding is consistent with the analysis by Nelson and his colleagues who found strong effects on educational domains from kindergarten to Grade 8. Also, narrative reviews by McLoyd (1998b) and Ramey & Landesman Ramey (1998) have argued that follow through programs with an educational component can build upon preschool education

(Nelson et al., 2003). What sets our finding apart from others is that the effects of follow through are present until the late adolescent period. Consistent with Nelson and his colleagues, we were also unable to determine how much more educational intervention would be required before a plateau on cognitive and educational impacts is reached.

#### **6.4: Program characteristics as moderators of outcome effect sizes**

Three program characteristics namely length, intensity, and number of program components were found to have positive effects on effect sizes during adolescent follow-up. Early childhood programs were disaggregated into three categories with respect to length; those equal to or less than 1 year, greater than 1 year but less than 3 years, and greater than 3 years. Insufficient data were available to compare programs that operated less than 1 year with other categories. Consequently, programs that operated for greater than 1 year but less than 3 years and greater than 3 years are compared.

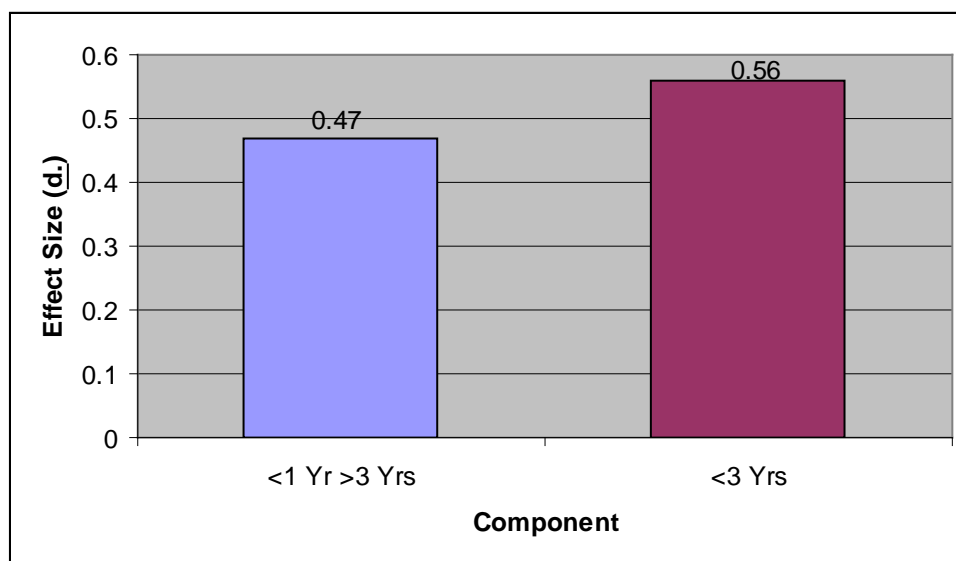
##### **6.4.1: Program length**

###### *A. ES Domain*

Early childhood programs whose duration was greater than 3 years demonstrated a larger average weighted mean on the ES domain ( $n = 5$ ;  $\underline{d} = 0.56$ ;  $\text{Var} = 0.007$ ;  $\text{CI Lower} = 0.387$ ;  $\text{CI Upper} = 0.726$ ;  $Q = 6.006$   $p < .001$ ) than programs whose duration was greater than 1 year but less than 3 years ( $n = 4$ ;  $\underline{d} = 0.47$ ;  $\text{Var} = 0.018$ ;  $\text{CI Lower} = 0.213$ ;  $\text{CI Upper} = 0.737$ ;  $Q = 2.577$   $p < .001$ ) (See Figure 6.4). The

range of (d), or unweighted mean, for programs whose duration was  $>1<3$  years was 0.0 – 0.94 (Mean = 0.406; SD = 0.343; Sample Var = 0.117), while the range of (d) for programs whose duration  $>3$  years was 0.24 – 0.82 (Mean = 0.56; SD = 0.239; Sample Var = 0.057).

**Figure 6.4: Average weighted effect sizes for the outcome educational success for program duration  $<1$  year  $>3$  years and  $<3$  years**



### *B. CD Domain*

The average weighted d for the CD outcome domain demonstrated that programs whose duration was greater than 3 years was smaller ( $n = 9$ ;  $d. = 0.32$ ;  $Var = 0.003$ ;  $CI\ Lower = 0.212$ ;  $CI\ Upper = 0.438$ ;  $Q = 6.283$   $p < .001$ ) than programs whose duration was  $>1<3$  years ( $n = 5$ ;  $d. = 0.38$ ;  $Var = 0.009$ ;  $CI\ Lower = 0.201$ ;  $CI\ Upper = 0.577$ ;  $Q = 1.144$ ,  $p < .001$ ). However, the range of (d), or unweighted mean, for programs whose duration was  $>1<3$  years was 0.16 – 0.51 (Mean = 0.40; SD = 0.147; Sample Var = 0.021), while the range of (d) for programs whose duration  $>3$  years was 0.18 – 0.74 (Mean = 0.43; SD = 0.178; Sample Var = 0.031).

### *C. SED Domain*

The average weighted  $\underline{d}$  for the SED outcome domain demonstrated that programs whose duration was greater than 3 years was smaller ( $n = 3$ ;  $d. = 0.12$ ;  $Var = 0.002$ ;  $CI\ Lower = 0.044$ ;  $CI\ Upper = 0.205$ ;  $Q = 1.622$   $p < .001$ ) than programs whose duration was  $>1 < 3$  years ( $n = 3$ ;  $d. = 0.44$ ;  $Var = 0.014$ ;  $CI\ Lower = 0.106$ ;  $CI\ Upper = 0.565$ ;  $Q = 2.433$ ,  $p < .001$ ). The range of ( $d$ ), or unweighted mean, for programs whose duration was  $>1 < 3$  years was  $0.0 - 0.65$  ( $Mean = 0.29$ ;  $SD = 0.329$ ;  $Sample\ Var = 0.108$ ), while the range of ( $d$ ) for programs whose duration  $>3$  years was  $0.11 - 0.40$  ( $Mean = 0.21$ ;  $SD = 0.165$ ;  $Sample\ Var = 0.027$ ).

### *D. CJ Domain*

The average weighted  $\underline{d}$  for the CJ outcome domain demonstrated that programs whose duration was greater than 3 years was smaller ( $n = 3$ ;  $d. = 0.22$ ;  $Var = 0.002$ ;  $CI\ Lower = 0.118$ ;  $CI\ Upper = 0.329$ ;  $Q = 2.050$   $p < .001$ ) than programs whose duration was  $>1 < 3$  years ( $n = 2$ ;  $d. = 0.37$ ;  $Var = 0.019$ ;  $CI\ Lower = 0.100$ ;  $CI\ Upper = 0.639$ ;  $Q = 0.105$ ,  $p < .001$ ). The range of ( $d$ ), or unweighted mean, for programs whose duration was  $>1 < 3$  years was  $0.32 - 0.41$  ( $Mean = 0.365$ ;  $SD = 0.063$ ;  $Sample\ Var = 0.004$ ), while the range of ( $d$ ) for programs whose duration  $>3$  years was  $0.20 - 0.48$  ( $Mean = 0.30$ ;  $SD = 0.159$ ;  $Sample\ Var = 0.025$ ).

### *E. FW Domain*

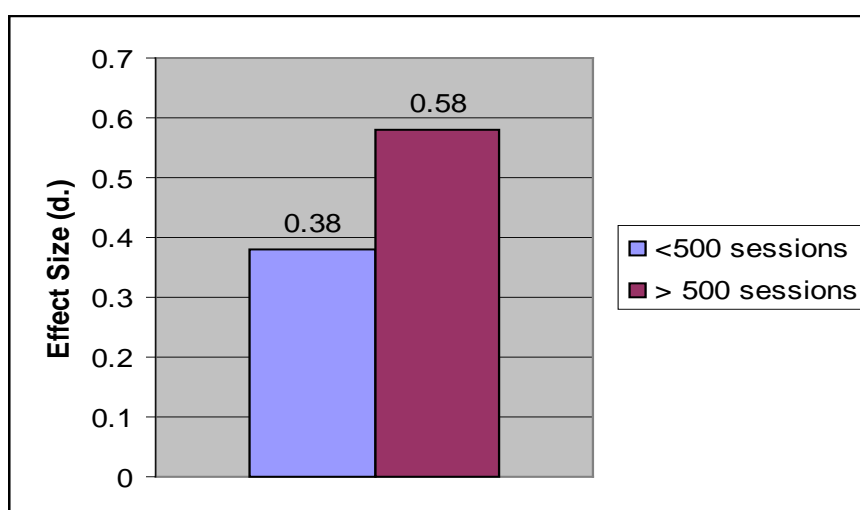
The average weighted  $d$  for the FW outcome domain demonstrated that programs whose duration was greater than 3 years was smaller ( $n = 3$ ;  $d = 0.14$ ;  $Var = 0.004$ ;  $CI\ Lower = 0.008$ ;  $CI\ Upper = 0.280$ ;  $Q = 2.056$   $p < .001$ ) than programs whose duration was  $>1 < 3$  years ( $n = 2$ ;  $d = 0.38$ ;  $Var = 0.014$ ;  $CI\ Lower = 0.146$ ;  $CI\ Upper = 0.614$ ;  $Q = 0$ ,  $p < .001$ ). However, the range of ( $d$ ), or unweighted mean, for programs whose duration was  $>1 < 3$  years was  $0.0 - 0.38$  (Mean =  $0.19$ ;  $SD = 0.268$ ; Sample  $Var = 0.072$ ), while the range of ( $d$ ) for programs whose duration  $> 3$  years was  $0.11 - 0.43$  (Mean =  $0.22$ ;  $SD = 0.185$ ; Sample  $Var = 0.034$ ).

Our analysis highlights that length of program (e.g.  $> 1$  year but  $< 3$  years; and  $> 3$  years) is an important moderator of success on the domain of educational success ( $> 1$  year but  $< 3$  years =  $.47$ ; and  $> 3$  years =  $.56$ ). However, we did find that results were not as impressive on other outcome domains; for example, cognitive development ( $> 1$  year but  $< 3$  years =  $.38$ ; and  $> 3$  years =  $.32$ ); social-emotional development ( $> 1$  year but  $< 3$  years =  $.44$ ; and  $> 3$  years =  $.12$ ); criminal justice outcomes ( $> 1$  year but  $< 3$  years =  $.37$ ; and  $> 3$  years =  $.22$ ); and family wellbeing ( $> 1$  year but  $< 3$  years =  $.44$ ; and  $> 3$  years =  $.14$ ). However, when we analysed the data taking into account ranges and sample means we found that programs whose duration was greater than 3 years demonstrated larger sample means than programs that were greater than 1 year but less than 3 years. Domains that demonstrated this effect include educational success (ES), cognitive development (CD), and family wellbeing (FW). Unfortunately, we had insufficient data to measure the effects of program length on the outcome domains social participation (SP) and deviancy (D).

### 6.4.2: Program intensity

Program intensity was coded according to the intended amount of sessions for both parents and children. Thus, intensity equalled the total number of sessions available to a participant. A full description of how we measured intensity is provided in Chapter 4. The average weighted  $d$  for the ES outcome domain demonstrated that programs whose intensity was greater than 500 sessions was larger ( $n = 7$ ;  $d. = 0.58$ ;  $Var = 0.007$ ;  $CI\ Lower = 0.417$ ;  $CI\ Upper = 0.744$ ;  $Q = 7.24$   $p < .001$ ) than programs whose intensity was  $\leq 500$  sessions ( $n = 4$ ;  $d. = 0.38$ ;  $Var = 0.021$ ;  $CI\ Lower = 0.099$ ;  $CI\ Upper = 0.672$ ;  $Q = 0.263$ ,  $p < .001$ ) (See Figure 6.5). The range of ( $d$ ), or unweighted mean for programs whose intensity was  $>500$  sessions was  $0.24 - 0.94$  (Mean =  $0.61$ ;  $SD = 0.261$ ; Sample  $Var = 0.068$ ), while the range of ( $d$ ) for programs whose intensity was  $\leq 500$  sessions was  $0.0 - 0.46$  (Mean =  $0.27$ ;  $SD = 0.195$ ; Sample  $Var = 0.038$ ).

**Figure 6.5: Average weighted effect sizes for the outcome educational success (ES) for programs whose intensity was  $<500$  sessions and  $>500$  sessions**



The average weighted  $\underline{d}$  for the CD outcome domain demonstrated that programs whose intensity was greater than 500 sessions was only slightly smaller ( $n = 9$ ;  $d. = 0.33$ ;  $Var = 0.003$ ;  $CI\ Lower = 0.219$ ;  $CI\ Upper = 0.445$ ;  $Q = 6.548$   $p < .001$ ) than programs whose intensity was  $\leq 500$  sessions ( $n = 5$ ;  $d. = 0.37$ ;  $Var = 0.009$ ;  $CI\ Lower = 0.178$ ;  $CI\ Upper = 0.559$ ;  $Q = 1.109$ ,  $p < .001$ ). However, the range of ( $d$ ), or unweighted mean, for programs whose intensity was  $>500$  sessions was  $0.18 - 0.74$  (Mean =  $0.46$ ;  $SD = 0.164$ ;  $Sample\ Var = 0.026$ ), while the range of ( $d$ ) for programs whose intensity was  $\leq 500$  sessions was  $0.16 - 0.51$  (Mean =  $0.35$ ;  $SD = 0.147$ ;  $Sample\ Var = 0.021$ ).

The average weighted  $\underline{d}$  for the SED outcome domain highlighted that programs whose intensity was greater than 500 sessions was smaller ( $n = 3$ ;  $d. = 0.12$ ;  $Var = 0.001$ ;  $CI\ Lower = 0.044$ ;  $CI\ Upper = 0.205$ ;  $Q = 1.622$   $p < .001$ ) than programs whose intensity was  $\leq 500$  sessions ( $n = 3$ ;  $d. = 0.34$ ;  $Var = 0.013$ ;  $CI\ Lower = 0.106$ ;  $CI\ Upper = 0.565$ ;  $Q = 2.433$ ,  $p < .001$ ). The range of ( $d$ ), or unweighted mean, for programs whose intensity was  $>500$  sessions was  $0.11 - 0.40$  (Mean =  $0.21$ ;  $SD = 0.164$ ;  $Sample\ Var = 0.027$ ), while the range of ( $d$ ) for programs whose intensity was  $\leq 500$  sessions was  $0.0 - 0.65$  (Mean =  $0.29$ ;  $SD = 0.329$ ;  $Sample\ Var = 0.108$ ).

The average weighted  $\underline{d}$  for the CJ outcome domain established that programs whose intensity was greater than 500 sessions was smaller ( $n = 3$ ;  $d. = 0.22$ ;  $Var = 0.002$ ;  $CI\ Lower = 0.118$ ;  $CI\ Upper = 0.329$ ;  $Q = 2.050$   $p < .001$ ) than programs whose intensity was  $\leq 500$  sessions ( $n = 2$ ;  $d. = 0.37$ ;  $Var = 0.019$ ;  $CI\ Lower = 0.100$ ;  $CI\ Upper = 0.639$ ;  $Q = 0.105$ ,  $p < .001$ ). The range of ( $d$ ), or unweighted mean for programs whose intensity was  $>500$  sessions was  $0.20 - 0.48$  (Mean =  $0.30$ ;  $SD =$



0.158; Sample Var = 0.025), while the range of (d) for programs whose intensity was  $\leq 500$  sessions was 0.32 – 0.41 (Mean = 0.37; SD = 0.063; Sample Var = 0.004).

Finally, the average weighted  $\underline{d}$  for the FW outcome domain demonstrated that programs whose intensity was greater than 500 sessions was smaller ( $n = 3$ ;  $d = 0.14$ ; Var = 0.004; CI Lower = 0.008; CI Upper = 0.280;  $Q = 2.056$   $p < .001$ ) than programs whose intensity was  $\leq 500$  sessions ( $n = 2$ ;  $d = 0.38$ ; Var = 0.014; CI Lower = 0.145; CI Upper = 0.614;  $Q = 0$ ,  $p < .001$ ). However, the range of (d) for programs whose intensity was  $>500$  sessions was 0.11 – 0.43 (Mean = 0.22; SD = 0.185; Sample Var = 0.034), while the range of (d) for programs whose intensity was  $\leq 500$  sessions was 0 – 0.38 (Mean = 0.19; SD = 0.269; Sample Var = 0.072).

Our examination of program intensity as a moderator of program success found that a program whose intensity was greater than 500 sessions demonstrated positive results on the outcome domain educational success ( $>500$  sessions = .58; and  $\leq 500$  sessions = .38). However, using weighted means we found either results to be in the favour of  $\leq 500$  sessions or no difference between either levels of intensity on other outcome domains. When we analysed the data taking into account ranges and sample means we found that programs whose intensity was greater than 500 sessions demonstrated larger sample means than program that were less than or equal to 500 sessions. Domains that demonstrated this effect include educational success, cognitive development, and family wellbeing. Once again, we had insufficient data to measure the effects of program length on the outcome domains social participation and deviancy.

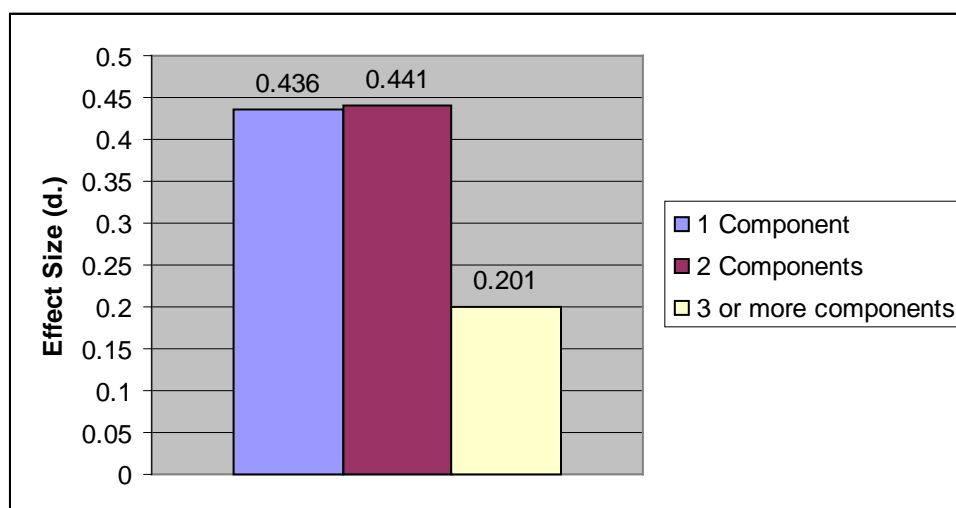
Combining length and intensity of programs we find that both are important moderators on some outcome domains (educational success (0.56), cognitive development (0.33), and deviancy (0.49)). Moreover, ranges of effect sizes indicate that programs that are longer and more intense have more effect on outcomes than those that are shorter and less intense. This finding is consistent with Nelson et al. (2003) and Ramey & Landesman Ramey (1998) who argue that that length and intensity of programs for children would be related to both child outcomes as well as parent-family outcomes.

#### **6.4.3: Number of program components**

In our analysis we measured whether number of program components (e.g. single component – structured preschool program; 2 components – home visitation plus structured preschool program etc.) was a potential moderator of program success on outcomes during adolescence. We found that the average weighted d for all outcomes based on number of program components demonstrated that single component programs were not as effective overall ( $n = 6$ ;  $d. = 0.436$ ;  $Var = 0.008$ ;  $CI\ Lower = 0.252$ ;  $CI\ Upper = 0.619$ ;  $Q = 2.827$   $p < .001$ ) compared to programs that had 2 components ( $n = 23$ ;  $d. = 0.441$ ;  $Var = 0.002$ ;  $CI\ Lower = 0.351$ ;  $CI\ Upper = 0.531$ ;  $Q = 9.82$ ,  $p < .001$ ). However, programs that were comprised of 3 or more components demonstrated a smaller weighted d when compared to programs that had 1 or 2 components ( $n = 17$ ;  $d. = 0.201$ ;  $Var = 0.0008$ ;  $CI\ Lower = 0.145$ ;  $CI\ Upper = 0.256$ ;  $Q = 28.05$ ,  $p < .001$ ) (see Figure 6.6). The range of (d), or unweighted mean, for all outcomes based on number of program components were as follows: 1 component 0.24 – 0.94 (Mean = 0.471; SD = 0.246; Sample Var = 0.060); 2 components 0.16 –

0.79 (Mean = 0.460; SD = 0.169; Sample Var = 0.028); and 3 or more components 0 – 0.82 (Mean = 0.259; SD = 0.217; Sample Var = 0.047).

**Figure 6.6: Average weighted effect sizes for number of program components**



Our analysis did not find evidence to support the assertion that number of program components is an important moderator of program success in the adolescent years. However, evidence demonstrates that when number of program components is measured during earlier life phases (e.g. early childhood) the outcome is significantly different. Moreover, results indicate that multi-component programs tend to have greater effect than single component programs particularly in treating or preventing conduct disorder (Foster, Olchowski, & Webster-Stratton, 2007). However, during the adolescent years we found only a small difference in overall effect size between single component programs (0.436) and 2 components (0.441). Our analysis also revealed that programs with three or more components had a significantly lower effect size (0.201) when compared to programs with one and two components. This sample was weighed down by the poor findings across all domains for studies 4 and 5. Further,

overall results were most probably biased due to the small number of longitudinal studies that follow-up early childhood intervention programs into the adolescent years.

Consequently, one should not assume (based on our findings during adolescence) that programs that have multiple components are any less favourable than programs with one or two components. As argued by Febbraro (1994), Reynolds et al., (2001), and Nelson et al. (2003), multi-component programs help families to meet a variety of needs, and without this network of support for families, outcomes beyond those directly affecting the child would be affected. This claim is well substantiated by the literature where it is argued that comprehensive, multi-component early childhood programs provide positive outcomes on children's social-emotional, educational, cognitive and family well-being (Nelson et al., 2001; Weissberg & Greenberg, 1998; Yoshikawa, 1994; Zigler et al., 1992). The meta-analysis performed by MacLeod and Nelson (2000), which focused specifically on early intervention programs designed to promote family wellness and prevent child maltreatment found effect sizes of approximately .58 for comprehensive, multi-component programs when compared to single component programs.

### **6.5: Fail-safe $n$**

Meta-analyses are vulnerable to bias (see Chapter 5, Section 5.1.3) given that studies that demonstrate non-significant findings are rarely published; this is known as the fail-safe  $n$  dilemma (Hedges & Olkin, 1985). The fail-safe  $n$  for  $n = 17$  studies and the overall weighted mean ( $\bar{d}$ ) is 0.35, which highlights that another 13 studies with non-significant findings would be required to reduce the overall weighted mean ( $\bar{d}$ ) to a small effect size of (0.20). It is argued that it is unlikely that there are 13 missing

longitudinal studies (with null results) of early childhood interventions with follow-up through adolescence. We base this conclusion on the argument that programs that met the criteria for inclusion into the meta-analysis were selected from approximately 5,200 abstracts. As stated in Chapter 5, we believe that the search uncovered all of the relevant studies identified in several narrative reviews of the literature

### **6.6: Limitations of our meta-analysis**

The first limitation of this study was the lack of longitudinal studies available to successfully measure the effect of early childhood intervention programs on the outcomes of deviancy, social participation and criminal justice and on the potential program moderators of length and intensity. Past meta-analyses have also reported the dearth of studies on long-term impacts (e.g. Farrington and Welsh, 2003; and Nelson et al., 2003). Only with future longitudinal investigations during adolescence and beyond will we be able to make more accurate judgements on the overall effectiveness of these programs on all outcome constructs.

The second limitation involves the grouping of outcome variables. Endemic to meta-analytic reviews is the “apples and oranges” problem. All meta-analyses are limited to using broad groupings of constructs (e.g. educational success, cognitive development, social-emotional development) to display outcome measures reported in the studies. The problem with this is that broad grouping tend to “gloss over” (Nelson et al., 2003, p.33) important differences among the various outcome indicators and their constructs. Once again, future longitudinal investigations should place more emphasis on fine-grained outcome constructs and their indicators. For example, indicators of deviancy should include: rates of delinquent behaviour; drug use (e.g.

marijuana and alcohol); lying about age (e.g. false Id); running away from home; caught breaking the law; and gang involvement. Indicators of social patterns include: casual employment in teen years; socio-economic success; engaged in skilled jobs (e.g. electrician); and adolescent response to problems (e.g. active response). Indicators of social responsibility include: rates of juvenile arrest; rates of violent and non-violent arrest; incarceration; petition requests to juvenile court; and adjudicated as a person in need of supervision due to incorrigible behaviour. With these detailed indicators it would be possible to place greater focus on how indicators affect the outcomes of the associated outcome constructs.

Finally, similar to the analysis performed by Nelson and his colleagues (2003), we were unable to tease apart some of the theoretical constructs such as strengths or empowerment orientation and cultural sensitivity of the intervention due to limited information on these constructs provided in the studies available for inclusion.

### **6.7: Implications of our meta-analysis**

Research findings asserting that early childhood intervention programs are good evidence-based policy are, based on our findings, quite legitimate claims. Not only are the results of this analysis and other past meta-analyses (e.g. Farrington & Welsh, 2003; MacLeod & Nelson, 2000; Nelson et al., 2001; Nelson et al., 2003) clear indications of the potential benefits, but they have also been demonstrated to be a good social investment from an economic perspective (see Greenwood et al., 2001; Karoly et al., 1998; Olds et al., 1993).

We concur with findings by Eckenrode et al. (2000a) and comments by Leventhal (2001) that home visitation programs reduce rates of child maltreatment and support the growth of social-emotional outcomes for both children and their mothers. We also agree with Nelson and his colleagues (2003) that early childhood programs with an educational component produce positive outcomes on children's cognitive development and educational success beyond the early years of primary school; and from our research we argue that this extends well beyond this period into late adolescence and early adulthood. Farrington and Welsh (2003) also argue that family-based crime prevention programs reduce rates of delinquency and anti-social behaviour in children and adolescents. Our views coincide with those by MacLeod and Nelson (2000), Nelson et al. (2001), Weissberg and Greenberg (1998), Yoshikawa (1994), and Zigler (1992) that multi-component early intervention programs do provide positive outcomes on children's social-emotional, educational, cognitive and family well-being. Our research also suggests that both length and intensity of programs are important moderators on some outcome domains (educational success, cognitive development, and family wellbeing) well into the adolescent years.

Gaps, however, still exist in our knowledge and understanding of developmental prevention and its effects into, and beyond the adolescent years. For example, the literature argues that more research is required to understand how children from various ethno-racial backgrounds differ. Ways must be found to customise programs to suit these variations in ethnicity to make them more beneficial to the target groups (Homel et al., 2006; The Developmental Crime Prevention Consortium, 1999). As argued by McLoyd (1998b) researchers also need to evaluate

the effectiveness of lengthy and intensive interventions on children from different ethno-racial backgrounds. Future research would also benefit from understanding the effects of early intervention, beyond the childhood years, on community-level outcomes.

## **6.8: Conclusion and the link to part 2 of the thesis**

Although early childhood intervention programs have been proven to be good evidence-based policy, results of the meta-analysis clearly demonstrate that deciding upon which program or group of programs to fund is a highly complex decision. Policy-makers face two dilemmas. First, policy-makers need to choose which program will produce the desired outcome/s at a given life phase (e.g. adolescence); and secondly, attempt to determine the economic costs and benefits of their decisions.

Regarding the first dilemma, policy-makers argue that their decisions are based on results from empirical research. This may be true; however, we argue that such complex decisions, incorporating multiple elements (e.g. goals, outcome constructs, outcome indicators etc.) cannot be made without an adequate tool for modelling complex problems under the veil of risk and uncertainty. In other words, we argue that policy decisions which are currently made with results from empirical research, but in an unstructured manner. Part 1 of this thesis attempted, for the first time, to provide an overall picture of the potential outcomes of early childhood intervention projects on at-risk children and their families during adolescence. It is important to note here that decisions should be made with findings from empirical research that is life phase specific. In other words, attempting to choose a program that potentially reduces rates of delinquent behaviour, improves rates of educational



success and reduces rates of children coming into contact with the criminal justice system using outcome data that is not specific to this life phase is a mistake. Rather, a systemic approach to aggregating the results of empirical research conducted during that life phase that focuses on evidence-based research is critical. However, it is argued that another step is required to make policy decisions more structured and reliable. Therefore, in Part 2 of the thesis, we attempt to develop a method for making complex decisions, using the results of evidence-based research related to achievement of a specific goal. Then in a structured and reliable manner, this provides a vehicle for making decisions (or policy options) in respect to attaining that goal under the veil of risk and uncertainty.

Policy makers have also attempted to make decisions incorporating various methods of determining the economic effectiveness of early childhood intervention program options/alternatives. For example, cost-benefit, cost-effectiveness and cost-savings studies have been conducted of many well known early childhood intervention projects (e.g. Perry Preschool, Elmira Nurse Home Visitation Program, and the Abecedarian Project) (see Chapter 1). However, when attempting to incorporate elements of utility into their analyses, they have fallen short. This is partly due to the inability of their methods to identify common metric outcomes that pertain to improvements in quality of life. As stated in Chapter 1, methods of attaining utilities or preference values for domains of quality of life are not new. Health economists, have with some success, gained preference values for potential outcomes that are condition specific (see Chapter 1). However, criminological research has not succeeded in reaching this point. Consequently, economic analyses are limited to cost-benefit, cost-savings, and cost-effectiveness analyses. Therefore, Study 2 also

provides a method for gaining the preference values of various early childhood intervention alternatives and their effect on non-health-related quality of life domains during the adolescent life phase. Preference values can then be used to determine the cost-utility of various early childhood intervention programs. It must be noted that without a meta-analysis of the effects of early childhood interventions on the adolescent life phase (Part 1 of the thesis), decisions regarding the potential outcomes and their perceived preference values would be based on limited evidence. It is with this information that individuals can make more informed decisions on the values that they attach to a given outcome and how they believe this will contribute to at-risk individuals non-health related quality of life.

The aim of Part 2 of the thesis is to provide a solution to two dilemmas faced by policy makers. As stated above, policy-makers need to choose which program will produce the desired outcome/s at a given life phase (e.g. adolescence); and secondly, attempt to determine the economic costs and benefits of their decisions. However, before demonstrating the use of our new methodology for resolving these dilemmas, an overview of the theoretical and philosophical foundations of economic analysis, together with a discussion regarding individual and social utilities is required. The aim of this discussion is to provide a foundation for subsequent chapters which identify various methods of attaining utilities/preference values and propose a method which can be incorporated to address the two dilemmas.

## **CHAPTER 7: ECONOMIC ANALYSIS: PREFERENCES AND UTILITIES**

As stated at the end of Chapter 6 (Section 6.8) the aim of Part 2 of the thesis is to provide potential solutions to two dilemmas faced by policy makers. The first dilemma regards the lack of a methodology to make well informed choices on resource allocation and structured decision-making with respect to alternative policy options for early childhood interventions. Secondly, the absence of a rigorous tool that permits identification of the economic impact of early childhood interventions on salient outcomes associated with non-health related quality of life (e.g. educational success, cognitive development, social-emotional development) throughout the life-course (e.g. early childhood and adolescence). However, before we introduce a methodology that provides solutions to both dilemmas, an examination of the theoretical and philosophical foundations of economic analysis, together with a discussion regarding individual and social utilities is conducted.

Chapter 7 (Section 7.1) begins by discussing the value-of-life literature highlighting the difficulty of valuing alternative life courses. In order to draw closer to developing a methodological standard that can value alternative life courses, we discuss the literature on cost-effectiveness (CEA) (Section 7.2) and cost-utility analysis (CUA) (Section 7.3). Section 7.4 examines the concept of ‘utility’, discussing the individual utility function, social utility functions, and Pareto Optimality. Section 7.5 explores methods for measuring preferences including: rating scales and their variants; standard gamble (health related); time trade-off (health related); and multi-attribute utility. Section 7.6 discusses the issue of choosing the best method, with the goal of providing a potential solution to the two dilemmas faced by policy makers, as

discussed in Chapter 5. Section 7.7 introduces the analytical hierarchy process method and preferences. Finally, Section 7.8 provides a summary of the chapter and briefly introduces Chapter 8.

### **7.1: Valuation of alternate life courses**

The value-of-life literature (Krupnick, Alberini, Cropper, Simon, O'Brien, Goeree, & Heintzelman, 2002; Viscusi, 1993, 2000) begins with the premise that life is a qualitatively different and superior state than death. Nevertheless, as we go about our everyday lives we make choices that reflect our attitude towards various life-threatening risks (Viscusi, 1993). Applying the method of 'willingness-to-pay' (e.g. the maximum price an individual is willing to pay for a particular good or service), economists are able to calculate the implicit price being paid in order to avoid or minimise life-threatening risks (e.g. willingness-to-pay for safety devices such as seat belts in cars or air bags), and determine the minimum amount of compensation a person is willing to accept if they take on these risks (e.g. compensating wage differentials for working in life-threatening or unpleasant work environments) (Coursey, Hovis, & Schulze, 1987; Nagin, 2001).

A major difficulty confronted when seeking to adapt the value-of-life methodology to the area of early childhood developmental intervention, is that in order to apply this methodology economists examine choices made explicitly by individuals (Nagin, 2001). For example, the choice between purchasing a car with particular safety features with verifiable impacts on risks of death in certain types of crashes with a car lacking these safety features but identical in all other characteristics. However, children and families who live in disadvantaged

communities often lack this capacity to make such explicit choices. This is particularly salient for children whose life course is largely beyond their control due to biological inheritance, parents' income, or the social and economic instability of the community in which they live (Golly et al., 1998; Homel et al., 2001). Moreover, a child's life course is further dependent upon the choices of others. For example, "...whether parents choose to invest their time and energy into building a child's personal capital rather than fulfilling their own needs" (Nagin, 2001, p.367).

The lack of capacity to make choices is not limited to children; it may also extend to a child's parents. For example, most parents would not deliberately choose a life of crime for themselves or their children. However, some parents find themselves in a position whereby their choices are limited, making "choices and compromises based on alternatives that they perceive before them" (Elder, Johnson, & Crosnoe, 2003, p.11). Therefore, it can be argued that most choices are bounded by the confines of an individual's life-course trajectory. Elder, Johnson and Crosnoe (2003) in partial agreement state: "...individuals construct their own life course through the choices and actions they take within the opportunities and constraints of history and social circumstance" (p. 11). That is, while recognising that there exist boundaries around an individual's life course that are difficult to break down, there is some scope for altering outcomes within this general bounds.

Nagin (2001) argues that adapting the value-of-life methodology to assess the effectiveness of developmental interventions represents a promising strategy; for example, one could focus on valuing a parent's investment (time and money) into their child's intellectual, social and moral development. However, he also recognises

that "...a difficult but still tractable analytical and statistical problem is to establish the relationship between these investments and the change in the child's life chances" (p.368). Alternatively, Nagin argues that one could survey parents to ascertain their willingness to financially contribute to their child's improved life chances, or their willingness to support, via increased taxes, a program aimed at all children. However, such an approach would reveal a disparity between "...the amount and quality of actual and projected parental investments" (Nagin, 2001, p.368). This highlights the commonly acknowledged difficulty with the willingness to pay methodology. Namely, that willingness to pay is income contingent, yet income levels in society are not equal. A stated willingness to pay of \$10 per week for someone earning \$100 per week is arguably placing a greater value on a particular program than a stated willingness to pay \$10 per week for someone earning \$1000 per week. Yet, the willingness to pay methodology would imply that the two individuals placed the same value on them. In addition, Nagin posits: "...the *raison d'être* for state intervention in the form of targeted child development programs such as Head Start is that some parents lack the capacity or will to make a minimally acceptable investment in their children "(p.369).

If the reason for state intervention in child development programs is based on parents' capacity, or the lack of capacity, to make an acceptable investment in their children, one may be led to question the minimum level of investment approach. To ascertain the level of investment considered a minimum, one could by contrast be guided by the literature on educational equality (Curran, 1995). Curran posits that:

...if one takes the central goal of education to lie beyond immediate results of instruction, to something like the broadening of life options or enhancement of

socio-economic status, then one will almost certainly speak of ‘equality of educational opportunity’ and take its substance to be something like equalization of life prospects of middle-class status, or more modestly and plausibly, equalization of opportunity to get an education that will improve these prospects...all children regardless of background, should have the right to this threshold of social inclusion (p.24).

Nagin (2001) argues that this standard should also be applied to developmental intervention services for at-risk children and their families.

To draw closer to developing a methodological standard that can value alternative life courses in the context of early childhood developmental intervention programs, one must move on to examine the literature on cost-effectiveness (CEA) and an adaptation of CEA, cost-utility analysis (CUA).

## **7.2: Cost-effectiveness analysis**

Cost-effectiveness analysis (CEA) typically refers to the evaluation of a program and its alternatives with respect to identifying the costs and effects of producing a given outcome (Levin & McEwan, 2001). For example, typical evaluations focus on a choice of a particular intervention and its alternatives with respect to a given objective (e.g. increasing academic test scores). Results of the program and its alternatives can then be assessed according to their effects on a given objective (e.g. comparison of test scores at 1 year follow-up). With comparable rates of effectiveness between the program and its alternatives and information related to their associated costs per participant, we are able to develop a cost-effectiveness (CE)

ratio for comparing the alternatives. Thus, cost divided by the difference in levels of effectiveness produces a CE ratio. Results are typically expressed as cost per unit of effect. Essentially, a C/E ratio is the incremental price of obtaining a unit of outcome from a given program/intervention when compared with its alternatives (Garber et al., 1996). The obtained ratios can then be used as a decision-oriented tool, whereby the most preferred alternative would be the one that demonstrates the lowest cost for a given increase in the objective (e.g. cost for each potential school drop-out averted). It is argued that choosing the most cost-effective alternative has the potential to maximise the efficiency in investment so that other areas can receive funding (e.g. other aspects of education). A major limitation to using this method as a means of measuring the economic effectiveness of early childhood intervention is that only programs with a similar or identical goal (or set of goals) can be compared. However, the benefits of most programs are usually much broader than this with outcomes often not directly comparable. Further, this method is typically limited to single variant rather than multi-variate analysis (Levin & McEwan, 2001). Consequently, when one attempts to measure multiple benefits across different outcome domains and compare programs that perform slightly different functions, one cannot adequately rank the alternatives.

If CEA is to be used as a tool for informing decision-makers on resource allocation, the effectiveness of such a tool is dependent upon the comparability and consistency of analyses of cost across a diverse range of alternatives. A brief examination of the literature reveals that investigators often make different assumptions about issues concerning costs to include, effects to measure, discount rate to be applied and the choice of in-kind costs to incorporate. Given that analyses tend



to use disparate methods, results across different analyses cannot be easily compared (Gold et al., 1996b).

To circumvent this problem, it is argued that a single fixed set of methodological standards should be adopted (Garber et al., 1996; Gold et al., 1996a; Gold et al., 1996b; Russell, Siegel, Daniels, Gold, Luce, & Mandelblatt, 2001). But how does one choose which method to adopt, given the myriad of choices available? Garber, Weinstein, Torrance and Kamlet (1996) propose that one examine the theoretical foundations of CEA. By doing this, one soon finds that a theory of decision-making can be either descriptive or normative. Garber et al. (1996) argue that since CEA is designed to be a practical tool for achieving societal goals, the theory should be normative.

Garber et al. (1996) argue that welfare economics may be considered the theoretical foundation for CEA. They state:

...welfare economics represents a comprehensive framework that provides answers to more methodological questions that arise in decisions from the 'societal perspective' than do any alternatives...Reliance on this theoretical framework can be based on its ability to inform specific issues from a societal perspective...Welfare economics provides guidance on such elements of CEA as how society should value resource costs and choose a discount rate for evaluation ( p.26).

However, this is not altogether infallible. If one fully adopts this method it does not adequately address "...which approach is best if one adopts a particular

decision-making perspective in which the constrained resources and the decision-maker's objectives are explicit"? ...Moreover, "they cannot directly answer questions that require reference to a fundamental set of values" (Garber et al., 1996, p.26). This highlights that there is some obvious disjuncture between societal value and optimisation techniques. Therefore, it is recommended that one move away from implicit welfare economics to a model that incorporates social goals. A method of moving beyond this limitation is cost-utility analysis.

### **7.3: Cost-utility (CU) analysis**

CU analysis is essentially a way of describing the relative preference strength of each outcome over a range of health and non-health related domains with respect to their associated costs. In other words, one evaluates program alternatives according to a comparison of their costs and perceived utility (Boardman et al., 2001). The method is similar to that of CEA in that a ratio is developed (costs divided by utility). The methods (CEA and CUA) are identical on the cost side but differ significantly on the effectiveness side. As stated above, levels of effectiveness are limited to programme-specific outcomes (e.g. increasing academic test scores) in CEA. However, in CUA, outcomes may be single or multiple, can be general as opposed to program specific and may incorporate a notion of value (Drummond et al., 1997). In CUA, individuals are asked to express their satisfaction with single or multiple measures of effectiveness. Once costs and utilities of the various outcomes are determined, one can choose the program alternative that provides the highest utility at the lowest cost (Levin & McEwan, 2001).

As discussed in Chapter 2, critics of current methods of conducting economic evaluations of early childhood intervention programs (e.g. cost-effectiveness, cost-benefit analysis and cost-savings analysis) do not target the more technical matters of economic analysis (e.g. choosing the right discount rate), but focus on the broader conceptual issues concerning the overall structure of the analysis. In short, these critiques highlight the need to develop a methodology that measures benefits across multiple domains, at different times, but still on an individual level. They also suggest that the dilemma faced by economists when attempting to place an economic value on early childhood (developmental) intervention has been in identifying a common metric that accurately values a qualitative improvement in an individual's life. As discussed earlier, the dispute surrounds the methods employed by previous cost-benefit analyses for summing benefits and costs. Nagin (2001) posit that the current approach is flawed because "...the difference between a socially and personally constructed life course and the destructive counterpart cannot be reduced to balancing outcomes such as lower special education, criminal justice, and welfare costs against outcomes such as higher earnings" (p.365).

We argued in Chapter 1 that CUA be considered the method of choice based on its ability to value not only the "intangibles" (as discussed by Nagin), but its ability to provide a more holistic approach to the measurement of developmental prevention program impact (e.g. increased public safety, or improved quality of life).

Exponents of CUA argue that its purpose is to measure the overall improvement in health and non health-related well-being/quality of life in individuals and society with the purpose of improving the distribution of resources to facilitate

such improvements (Levin & McEwan, 2001; Richardson, 1990; Richardson, Olsen, Hawthorne, Mortimer, & Smith, 1998). However, a brief exploration of the theoretical framework reveals that the method might need modification to remain a valid decision-making tool. As CUA is a tool for improving general welfare, based on the desirability of alternative allocations of resources from a societal point of view, it seems its context is squarely within welfare economics. However, as argued by Arrow (1963), the central problem with welfare economics is that “achievement of social maximisation is derived from individual desires” (p.941). Garber et al. (1996) posit that welfare economics is based on the assumptions that, first, individuals maximise a well-defined *preference function*. In other words, [their] “...utility or sense of well-being depends on, among other things, material consumption, and the utility or preference function follows certain conditions of rationality and logical consistency, [and secondly] ...that the overall welfare of society is a function of these individual preferences” (p.29). Given that much of the literature is based on the improvement of the welfare of an affected population, it is necessary to first measure individual well-being and then aggregate this to reach well-being at a societal level (Drummond, Stoddart, & Torrance, 1987; Garber et al., 1996). However, for the purposes of this study, we are interested in individual utility and the measurement of preferences with the goal of adapting cost-utility approaches to measure the utility that children and their parents gain as a result of early childhood intervention programs. However, before we proceed to methods of measuring preferences we need to discuss briefly utility theory including both individual utility and derivation of a social utility function.

## 7.4: Utility Theory

The term “utility” has, in recent times, become synonymous with the term “preference”. Because the words have been used interchangeably, with little distinction made between the two, much confusion arises when attempting to operationalise their meaning (Sen, 1991). However, the two terms are quite different. Drummond et al (1997) provide a simple but accurate distinction; they state: “Preference is the umbrella term that describes the overall concept; utilities and values are different types of preferences (p.146)”. Therefore, “utility and “preference” become separate entities when “...approaches are developed to define the concept more precisely and especially when attempts are made to measure it” (Drummond et al., 1997, p.146). Moreover, the term preference is often used to describe an individual’s choice between two alternatives and whether they prefer one alternative to the other (more or less) or they are indifferent between the alternatives. Isard and Smith (1982) state:

Given the actions of all other participants and  $z$ ...the outcome function  $\dots o^j$   
 $= f^j(a, z)$   $J = A, \dots, U$ ...associates an outcome  $o^j$  with each action  $a^j$  that  $J$  may take. The unit  $J$  is now taken to be able to state preferences among the outcomes associated with all his possible actions...we assume that the behaving unit is able to state, for every pair of outcomes  $o_a$  and  $o_b$ , whether he prefers  $o_a$  to  $o_b$ , whether he prefers  $o_b$  to  $o_a$ , or whether he is indifferent between the two. Also, for any three outcomes  $o_a$ ,  $o_b$ , and  $o_c$ , we assume that if he prefers  $o_a$  to  $o_b$  and  $o_b$  to  $o_c$ , then he prefers  $o_a$  to  $o_c$  (pp.18-19).

This final sentence in this quote incorporates the notion of transitivity, which is discussed in detail in Chapter 8 (Section 8.2.3).

When one considers preferences for different outcomes, it is desirable to assign numbers to represent these preferences. Consequently, the term utility is adopted whereby one assigns numbers on a scale to represent choices of preferences. Essentially, these numbers indicate how much utility (including zero) a preference derives from each outcome, thereby representing a cardinal preference function, or a cardinal utility function (Isard & Smith, 1982). The limitation with cardinal utility is its application in eliciting preferences for outcomes that are not numeric in nature. Hence, preferences for outcomes associated with non-health related quality of life cannot be represented using cardinal utility functions. Moreover, if one can use numbers, they may only represent whether an individual prefers one outcome to another. That is, the numbers used indicate no more than a set of rankings, whereby the highest number indicates the most preferred outcome and the lowest number the least preferred outcome. Isard and Smith (1982) posit that the ratios of these numbers have no meaning, nor do their absolute differences. Thereby, the numbers represent an ordinal preference function, often designated as an ordinal utility function. Between these two extreme situations is a concept known as relative utility, whereby ratios of the numbers associated with outcomes are meaningful but the numbers themselves are not (Isard & Smith, 1982). For example, Isard and Smith (1982) state: “a participant may be able to say that he values (prefers) one outcome twice as much as another but that it does not matter to him whether the first outcome has a number 200 and the second 100, or the first 150 and the second 75”(p.19). In this thesis we use the term preference to mean relative utility. In Section 7.7 we introduce the analytical hierarchy method (AHP) (Saaty, 1980) which provides pair-wise comparisons regarding the relative strength of preferences. The AHP method elicits preferences from people on a relative scale from 0 to 9, whereby Saaty uses mathematical

operations treating a pair-wise comparison matrix as if they are cardinal numbers. However, provided the vector of priorities and preferences are interpreted as relative preferences, the use of cardinal number operations to derive the vector of priorities is mathematically sound. Appendix C provides the axiomatic foundations of Sati's AHP process, providing mathematical support for the use of cardinal operations to derive vectors of priorities and implicit preference regarding the relative strength of one alternative over another.

Von Neumann and Morgenstern (Von Neumann & Morgenstern, 1944) published their theory of rational decision making under uncertainty, now called the Von Neumann-Morgenstern utility theory or expected utility theory. Based on a normative model, that is, how a rational individual should make decisions when faced with uncertainty, they defined what is meant by rational behaviour under uncertainty via a set of fundamental axioms. Bell and Fuqua (1986) present the axioms of expected utility theory as follows:

*Preferences exist and are transitive:* For any pair of risky prospects  $y$  and  $y^1$  either  $y$  is preferred to  $y^1$ ,  $y^1$  is preferred to  $y$  or the individual is indifferent between  $y$  and  $y^1$ . Additionally, for any three risky prospects  $y$ ,  $y^1$  and  $y^2$ , if  $y$  is preferred to  $y^1$ , and  $y^1$  is preferred to  $y^2$ , then  $y$  is preferred to  $y^2$ .

*Independence:* An individual should be indifferent between a two-stage risky prospect and its probabilistically equivalent one-stage counterparts derived using the ordinary laws of probability.

*Continuity of preferences:* If there are three outcomes such that  $x^1$  is preferred to  $x^2$ , which is preferred to  $x^3$ , there is some probability  $p$  at which the individual is indifferent between outcome  $x^2$ , with certainty or receiving the

risky prospect made up of outcome  $x^l$  with probability  $p$  and outcome  $x^3$ , with probability  $1 - p$  (Cited in Drummond et al., 1997, p.144).

It should be noted that utility theory (incorporating Von Neumann and Morgenstern's axioms) does not describe how individuals make decisions under risk and uncertainty. Rather, they are a prescriptive or normative model of how rational individuals should make decisions as defined by the basic axioms (Nease, 1996; Von Neumann & Morgenstern, 1944). Drummond et al., (1997) states: "Normative models are used to define approaches that individuals should take to be consistent with underlying theories; behavioral models are used to describe actual behaviour as found in reality"(p.146).

Vigorous debate regarding the axioms of von Neumann-Morgenstern utility theory has taken place since its inception. Many variations and alternatives have been proposed (see Allais, 1991; Cohen, 1996; Currim & Sarin, 1992; Kleindorfer, Kunreuther, & Schoemaker, 1993; Nease, 1996; Schoemaker, 1992; Tversky & Kahneman, 1992). To date, however, none of the proposed alternatives have taken over the position of the dominant normative theory. Moreover, von Neumann and Morgenstern's axioms have been widely applied in the business, government and health care sectors, forming and remaining the foundation for modern decision theory (Drummond et al., 1997).

#### **7.4.1: Individual utility functions**

Individual utility maximisation requires recognising individual preferences. These are represented by individual utility functions, "which relate their well-being to



their levels of consumption of a number of goods and services” (Garber et al., 1996, p.30). The problem with simple economic models is their reliance on assuming a world of complete certainty, where prices are known, random events do not occur and information is freely available. In practice, however, very little is certain. Most of our knowledge is imperfect and relatively uncertain (Arrow, 1963). As a result of this uncertainty, modelling preferences adopting CEA and CUA approaches (as used in both health and welfare economics) requires the use of expected utility theory (Dasgupta, 1995). According to expected utility theory, the decision-maker chooses between uncertain and sometimes risky actions by comparing their expected utilities. Probabilities are then attached to the various outcomes (Hirshleifer & Riley, 1992; Mongin, 1997). Garber et al (1996) state:

Quantitative representations of preferences, or utilities, are assigned to each possible outcome (e.g., health state) that may occur. To choose the best action, the probability of each outcome is multiplied by the utility of that outcome; these products are then summed across all possible outcomes in order to derive the expected value of utility (Garber et al., 1996, p.30).

In order to maximise one’s utility, one may be willing to give up some proportion of, for example, health care for exchange of another item of commodity, such as cash or, alternatively, give up cash for more health. In welfare economics, this would be considered a trading opportunity, in which both individuals are made better off. Such a trade, however, would result in less aggregate health for society as a whole (Garber et al., 1996). Garber argues: “Moreover, it would leave the poor people in worse health than the rich, although they would consider themselves better off than under the initial state of affairs” (Garber et al., 1996, p.32). Is this a socially and

ethically acceptable trade? It is, according to neoclassical welfare economics, as at the end of the day all parties have gained and consider themselves better off. The ‘extra-welfarist’ economist, however, would not consider this to be acceptable, given that the societal value of ‘good’ is something which cannot be traded, regardless of individuals’ willingness to pay for it (Garber et al., 1996; Williams, 1993).

#### **7.4.2: Social utility function**

The next problem faced by economists is how to aggregate individual utility to reach a social utility function. The social utility function can be defined as the aggregate of individual utilities; “...economists view the maximization of the social utility function as the ultimate goal of any resource allocation scheme” (Garber et al., 1996, p.32). *Strict utilitarianism* is one approach to aggregating social utility. This approach generally sums the utilities of individuals within society whilst allowing for different people to receive different weights based on their individual circumstances in society. For example, a person who is in poor health or who is impoverished might be given a greater weight compared to those who are not. However, a scan of the literature reveals that *strict utilitarianism* is an approach most economists avoid when measuring well-being, due mainly to the difficulty of selecting a specific weighting scheme that is fair and representative of all in society (Sen, 1995). Given the lack of consensus on how individual utilities can be aggregated to form a social utility function, one must turn to mainstream microeconomic theory to establish the effect of resource reallocation on social welfare.

### 7.4.3: Pareto optimality

When one attempts to measure the effect of a resource allocation on social welfare, it is prudent to begin with the microeconomic concept of *Pareto optimality*. This concept has been a useful guide to test the effects of resource allocation on social welfare. According to Garber et al. (1996);

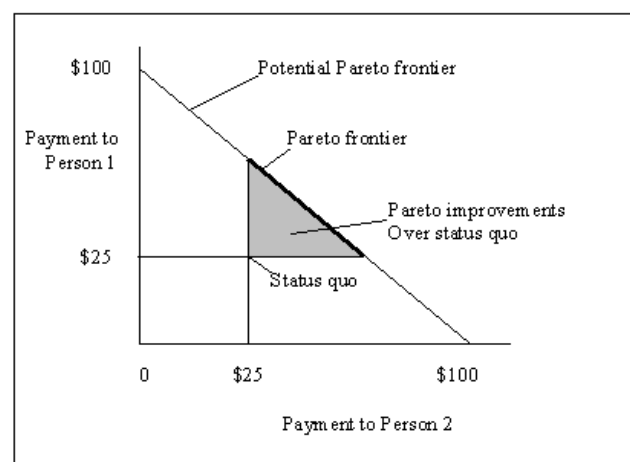
A resource distribution is considered to be Pareto-optimal when any change in the distribution must make someone worse off, even if others are better off...If the reallocation makes at least one person better off, and no one worse off, it is said to represent a Pareto improvement...Thus, when the effects of a change in policy or prices on individual utilities are known, but the specific social welfare function is not, the Pareto criterion can be used to test whether social welfare is improved (p.33).

The problem with this concept is that in the real world few policies produce only winners. For example, a federally-subsidised intervention program is generally allocated resources from taxation. "...typically, funds must be raised by taxes or another mechanism that imposes costs on some people that exceeds the benefits they can expect to receive"(Garber et al., 1996, p.33). As a consequence of the limitations of *Pareto optimality*, a less restrictive concept has been developed to measure resource allocation in which there are both winners and losers.

The Kaldor-Hicks criterion, or the compensation test, is described by Little (1950) as a less restrictive concept, whereby an evaluation situation considers both winners and losers from a reallocation of resources. The conceptual basis of cost-

benefit analysis, this compensation test considers a program to be welfare-enhancing if an allocation of resources makes at least one person better off without making anyone else worse off. The relationship between net social benefit and *Pareto efficiency* is described by Boardman et al. (Boardman et al., 1996): “If a policy has positive net benefits, then it is possible to find a set of transfers or side payments, that makes at least one person better off without making anyone else worse off” (p.30). Figure 7.1 illustrates this concept. In this example, the shaded triangle represents all the alternative allocations that would make at least one of the persons better off than the status quo without making the other worse off (Boardman et al., 2001, p.27).

**Figure 7.1: An illustration of Pareto efficiency**



Source: Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2001). *Cost-benefit analysis: Concepts and practice* (2nd ed.). Upper Saddle River, New Jersey: Prentice Hall.

Intuitively, the Kaldor-Hicks test is tantamount to the willingness of potential gainers to pay the maximum amount of money they consider the program to be worth, and losers to receive the minimum amount of money that would compensate them for their losses. If there is more money ‘in the hat’ at the end of the day, the allocation is considered to be a potential *Pareto improvement* (Boardman et al., 1996, 2001;

Garber et al., 1996). The problem with this approach is that the allocation between winners and losers may not occur. Consequently, "...the desirability of a program from the societal perspective cannot be determined without reference to the distribution of welfare, and a well-defined way of combining the welfare of different people into a social welfare function must again be invoked"(Garber et al., 1996, p.34).

Welfare economics assists in providing a derivative of the cost-effectiveness approach in order to elucidate decisions based on cost-effectiveness ratios, ensuring their equivalence with the Kaldor-Hicks criterion. The Kaldor-Hicks criterion states "...a policy should be adopted if and only if those who gain could fully compensate those who will lose and still be better off...as long as benefits are positive, it is at least possible that losers could be compensated so that the policy potentially could be Pareto improving"(Boardman et al., 2001, pp.29-30).

### **7.5: Methods of measuring preferences**

As discussed, non-economists tend to use the terms 'utility', 'value' and 'preference' interchangeably. However, they are quite distinct terms with respect to not only their concepts, but also how they are measured. Drummond et al., (1997) and Levin and McEwan (2001) describe two key aspects of the measurement process. The first involves the framing of the question; for example, are the outcomes in the question certain or uncertain. The second involves the way in which the subject is asked to respond; for example, is the subject asked to perform a scaling based task or to make a choice using introspection.

Questions which are framed under certainty (commonly termed the direct method) involve asking a subject to compare two or more outcomes (ensuring that the subject assumes that the outcome/s will occur with certainty) and choose between them on a scale. Levin and McEwan (2001) state: “When preferences are elicited by variants of this method, they are said to reflect the underlying value of outcomes, rather than utility” (p.200). By contrast, a question framed under uncertainty (commonly termed the indirect method or variable probability method) requires a subject to compare two alternatives in which at least one outcome is uncertain (Cohen, 1996). In this case, Levin and McEwan argue that preferences elicited by this method should be referred to as utilities.

Why do we need to make this distinction? There are two reasons. First, the variable probability method is based on the seminal theory of von Neumann and Morgenstern (1944) which provides a prescriptive or normative model of how rational individuals should make decisions as defined by the basic axioms. Secondly, utilities which are elicited by the variable probability method are said to rate an individual’s attitude toward risk (Von Neumann & Morgenstern, 1944).

Before individual methods of measuring preferences are discussed, we provide a basic interpretation of the methods, which clearly highlights the fact that what you get depends on your measurement tool (see Table 7.1). Table 7.1 highlights the methods used to elicit utility values from perceived preferences. It is argued by most economists (e.g. Spilker 1996b., Torrance, 1986) that the method in cell 4 is the only method which can truly capture the element of risk. Risk is measured by giving an individual a choice between two alternatives. Alternative 1 represents two possible

outcomes; e.g. patient is returned to normal health and lives for an additional  $t$  years (probability  $p$ ), or immediate death (probability  $(1-p)$ ). Alternative 2 contains an outcome of a chronic state  $i$  for life ( $t$  years). Probability  $p$  is altered until the respondent is indifferent between the two alternatives, where the preference value for state  $i=p$ . The difference between cells 1 and 3 highlights the difference between the concept of choosing and scaling. Moreover, cell 2 is empty as one cannot measure utilities (or measure risk) using scaling methods.

**Table 7.1: Methods of measuring preferences and utilities**

Response method	Question framing	
	Certainty (values)	Uncertainty (utilities)
Scaling	<div>1</div> Rating scale Category scaling Visual analogue scale Ratio scale	<div>2</div>
Choice	<div>3</div> Time trade-off Paired comparison Equivalence Person trade-off	<div>4</div> Standard gamble

Source: Drummond, M. F., O'Brien, B., Stoddart, G. L., & Torrance, G. W. (1997). *Methods for economic evaluation of health care programmes* (2nd ed.). New York: Oxford University Press (pp.146).

There are three widely used methods for measuring utilities; rating scales and their variants (cell 1 of Table 7.1), time trade-off and its variants (cell 3 of Table 7.1), and standard gamble methods (cell 4 of Table 7.1). The literature describing the various methods is vast (see Boardman et al., 1996, 2001; Drummond et al., 1997; Gold et al., 1996a; Levin & McEwan, 2001; Spilker, 1996b; Torrance, 1986;

Torrance, Siegel, & Luce, 1996c). In our discussion we include descriptions of both chronic and temporary states which incorporate both health and non health-related states of quality of life.

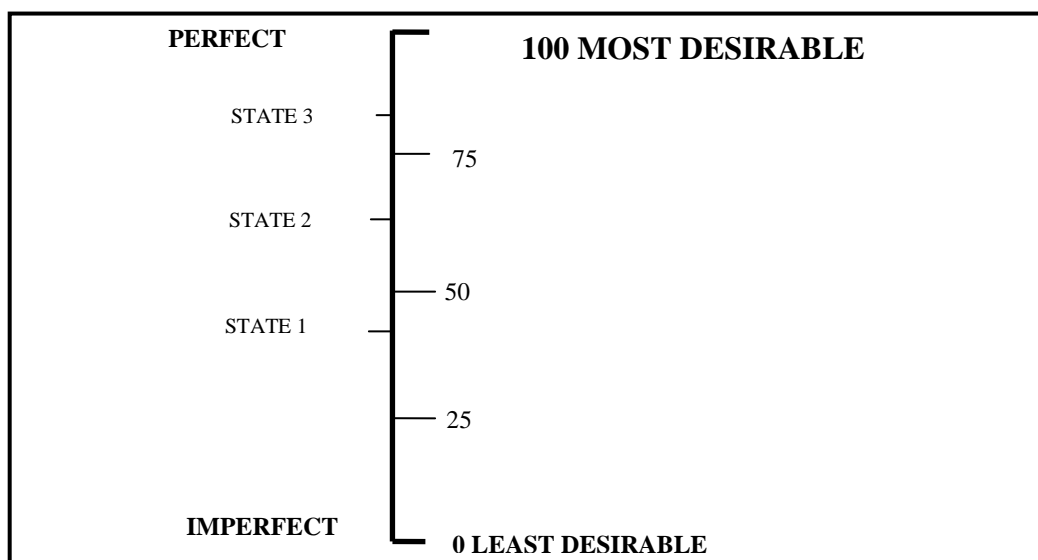
### **7.5.1: Rating scales and their variants**

The use of a rating scale involves the following steps:

1. A number of health/non-health related state scenarios are described which may include individual respondents subjective opinion of their current health/non-health related state, along with a series of hypothetical states that are condition specific. Scenarios may cover many dimensions such as mobility, pain and suffering, discomfort, cognition, social-emotional state, etc.
2. Respondents are then prompted to rank the various states along a visual analogue scale from 0 to 100 (sometimes called a feeling thermometer) (Figure 7.2).
3. The scores obtained provide a firm indication of ordinal rankings of the various states (or outcomes) and the intensity of the preferences.

Preferences for chronic states can also be measured on a rating scale. Chronic states must be described as irreversible and subjects must be provided with age of onset and age of death. These parameters should remain constant for all states that are measured relative to each other in a batch. As stated above, the rating scale is measured between 0 and 100; with the least desirable state (normally death) being 0 and the most preferred state (perfect health) being 100.



**Figure 7.2: The visual analogue scale**

Adapted from: Grabowski, H. G., & Hansen, R. W. (1996). Economic scales and tests. In B. Spilker (Ed.), *Quality of life and Pharmacoeconomics in clinical trials* (pp. 79-95). Philadelphia, New York: Lippincott-Raven.

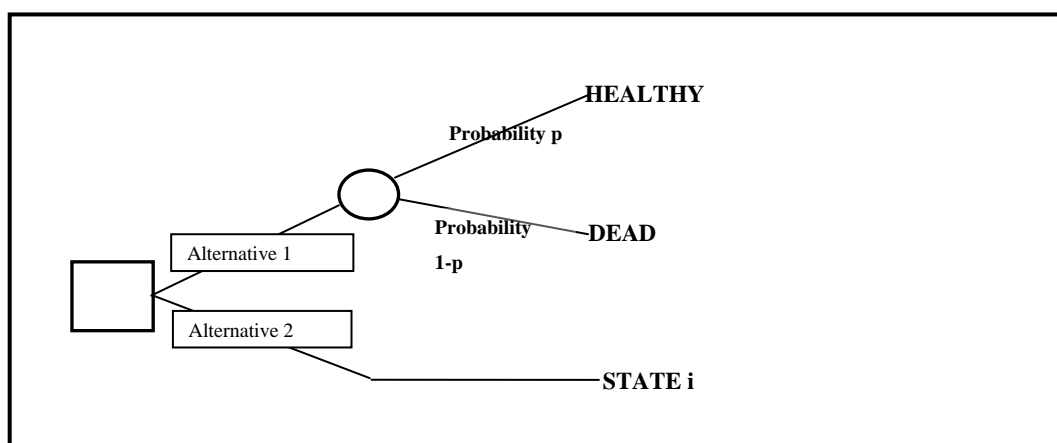
Preferences for temporary states can also be measured using a rating scale. States must be described as lasting for a specified period of time at which point one returns to a state of normality. As with chronic states, parameters should remain constant for all states that are measured relative to each other in a batch. Transforming to a standard 0-100 preference scale can be achieved by redefining the worst temporary state as a chronic state, of similar duration, and measuring its preference value with the same technique as used in measuring the chronic states. Values for other temporary states can be transformed by using a positive linear transformation (Drummond et al., 1997; Grabowski & Hansen, 1996).

### 7.6.2: Standard gamble (health related)

The use of the standard gamble approach when measuring preferences for chronic states involves the following steps:

1. Subjects are offered a choice between two alternatives. Alternative 1 represents two possible outcomes; patient is returned to normal health and lives for an additional  $t$  years (probability  $p$ ), or immediate death (probability  $(1-p)$ ). Alternative 2 contains an outcome of a chronic state  $i$  for life ( $t$  years).
2. Probability  $p$  is altered until the respondent is indifferent between the two alternatives, where the preference value for state  $i=p$  (Figure 7.3). Given that some subjects have difficulty relating to probabilities the method can be supplemented with visual aids such as a probability wheel (Torrance, 1986).

**Figure 7.3: Standard gamble for chronic health state preferred to death**

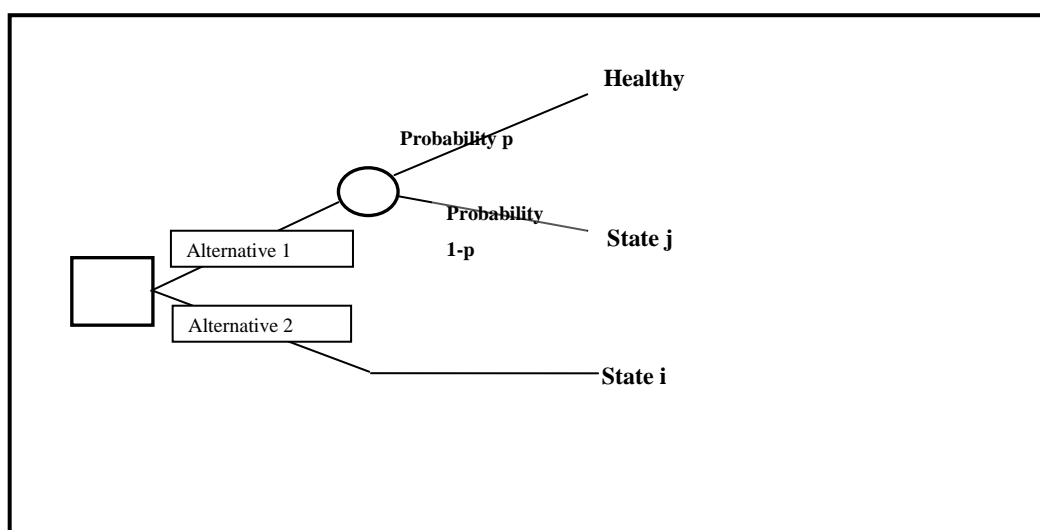


Source: Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5(1), 1-30.

Variations of this method are possible; this can be helpful when the two alternatives offered do not relate to healthy and dead. So long as the two alternative outcomes are such that one is preferred to the outcome of the certain alternative and the other is less preferred (dis-preffered) to it, the method can be adopted to relate the preference values of the three alternatives (Torrance, 1986).

Preferences for temporary states can also be measured by using the standard gamble method (Figure 7.4). In this case, intermediate states (i) are measured relative to the best temporary state (healthy) and the worst temporary state (temporary state j). Drummond et al. (1997) posit: "...all states last for the same duration, say t, followed by a common state, usually healthy. In this format the formula for the utility of state i for time t is  $h_i = p + (1-p) h_j$ , where i is the state being measured and j is the worst state" (p.154).

**Figure 7.4: Standard gamble for temporary health state**



Source: Drummond, M. F., O'Brien, B., Stoddart, G. L., & Torrance, G. W. (1997). *Methods for economic evaluation of health care programmes* (2nd ed.). New York: Oxford University Press.

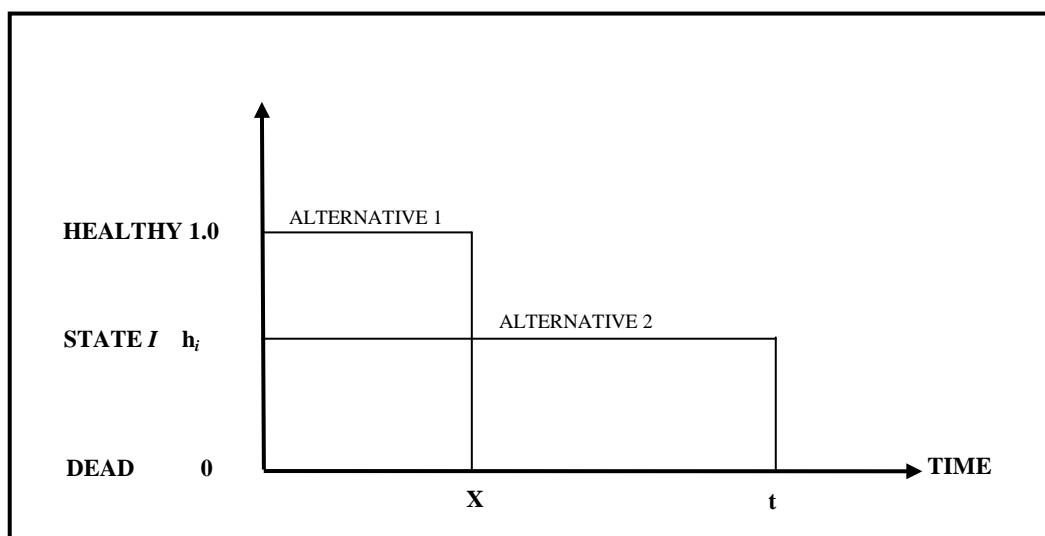
### 7.6.3: Time trade-off (health related)

Given the difficulty that some respondents have relating to probabilities, the time trade-off technique was developed. The technique was found to produce similar results to standard gamble, although its validation is limited to states preferred to death (Torrance, 1976). Measuring preferences for chronic states using the time trade-off approach involves the following steps:

1. Subject is offered the option of living for  $t$  years in perfect health or  $t$  years in another state that is less desirable (generally the state the researcher is attempting to score).
2. Perfect health,  $x$ , is then reduced systematically until the subject is indifferent between the shorter period in perfect health and the longer period in less desirable state.
3. The time trade-off score for the given state then equals  $x/t$ .

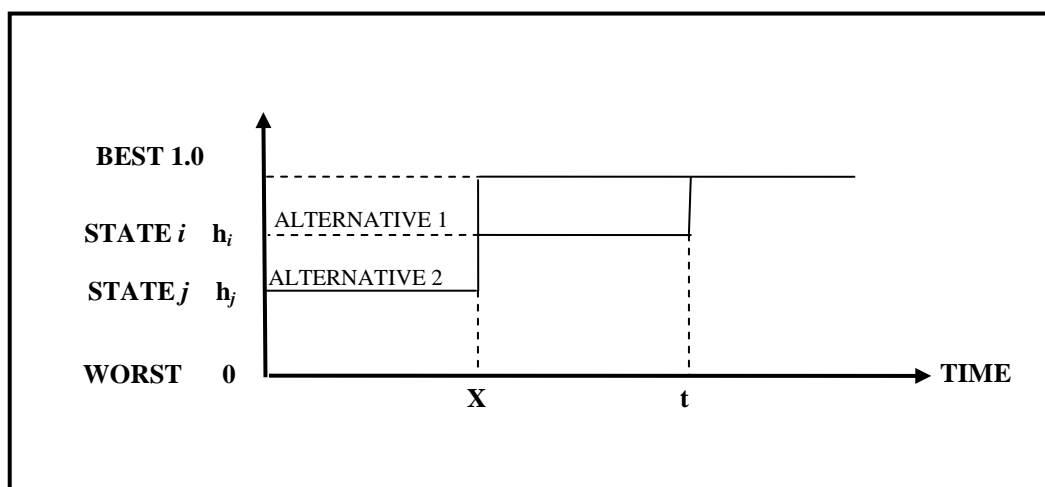
Figure 7.5 demonstrates the above steps, where an individual is offered two alternatives – alternative 1: state  $I$  for time  $t$  (life expectancy of an individual of an individual with a chronic condition) followed by death; and alternative 2 – healthy for time  $x < t$  followed by death;  $x$  is varied until the respondent is indifferent between the two alternatives, whereby  $h_i = x/t$  (Drummond et al., 1997; Torrance, 1986).

**Figure 7.5: Time trade-off for a chronic health state preferred to death**



Source: Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5(1), 1-30.

Preferences for temporary states may also be measured by using the time-trade-off method (see Figure 7.6). In this example, intermediate states  $i$ , are measured relative to the best state and the worst temporary state. Two alternatives are offered to the respondent. Alternative 1: temporary state  $i$  for time  $t$  (duration specified for the temporary state), followed by state of normality; and, Alternative 2: temporary state  $j$  for time  $x < t$ , preceded by state of normality. Time  $x$  is varied until the respondent becomes indifferent between the two alternatives. At this point, the preference state for  $i$  is  $h_i = 1 - (1 - h_j)x/t$  (Torrance, 1976). Similar to the rating scale and the standard gamble, if preferences for temporary states are to be transformed to a scale (0-1), the least preferred temporary state, or worst temporary state, must be re-defined as a short duration chronic state and measured using the method for measuring chronic states (Feeny, Torrance, & Labelle, 1996).

**Figure 7.6: Time trade-off for a temporary state**

Source: Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5(1), 1-30.

#### 7.5.4: Multi-attribute utility

The measurement techniques described above are used to measure preferences relating to single states. However, selecting specific preferences to be measured and combining them to represent a mathematical model of a subject's utility structure requires the application of a more sophisticated approach. This approach determines the utility values for all possible states in a given classification system (Drummond et al., 1997). To achieve this, a multi-attribute utility model must be adopted. It should be noted that the multi-attribute approach does not require the development of new methods of measuring preferences. Keeney and Raiffa (1976) extended the traditional von Neumann-Morgenstern utility theory to cover multi-attribute outcomes. Achieving this required an additional assumption namely utility independence among attributes.

Measuring preferences for outcomes (e.g. health outcomes) can be a very difficult and time consuming task. Consequently, attractive alternatives have been developed which potentially bypass the measurement task. They do this by using pre-scored multi-attribute health state classification systems that already exist. Examples include the quality of well-being scale (QWB), EuroQol (EQ-5D), and the Health Utilities Index (HUI). QWB classifies patients according to four health attributes; mobility, physical activity, social activity; and, symptom-problem complex. Scoring is based on category scale measurements (preference values) conducted on a random sample of the public (Kaplan & Anderson, 1996; Kaplan, Anderson, Wu, Mathews, Kozin, & Orenstein, 1989). EuroQol (EQ-5D) classifies patients according to six attributes; mobility, self-care, main activity, social relationships, pain, and mood. A subsequent revision of the system has reduced the number of attributes to five (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression). Each attribute consists of three levels; no problem, some problem, and major problems. Preferences are elicited by using the time trade-off technique on a random sample of the population (see Dolan & Gudex, 1995; Dolan, Gudex, Kind, & Williams, 1996). Three versions of the Health Utilities Index (HUI) exist (HUI1, HUI2, and HUI3). HUI1 consists of four attributes; physical functioning including mobility and physical activity, role function including self-care and role-activity, social-emotional function including emotional well-being and social activity, and health problem. Preference scores are measured using the time trade-off technique on a random sample of the population. The HUI2 system consists of six attributes; sensory and communications ability, happiness, self-care ability, pain or discomfort, learning and school ability, and physical activity ability. Preferences are measured using both a visual analogue scale and the standard gamble technique on a random

sample of parents of school children. HUI2 has also been modified for adult applications (Cadman, Goldsmith, Torrance, Boyle, & Furlong, 1986; Feeny, Boyle, & Torrance, 1995; Torrance, Feeny, Furlong, Barr, Zhang, & Wang, 1996a; Torrance, Furlong, Feeny, & Boyle, 1995). The HUI3 classification system is similar to the HUI2 classification system; however, the attribute 'fertility' is excluded. The sensory attribute is expanded to incorporate three dimensions: vision, hearing, and speech. Other changes are made to increase structural independence of the attributes. Preference scores are measured using a visual analogue scale and the standard gamble technique on a random sample of the population (Drummond et al., 1997).

## **7.6: Choosing the best method**

Although rating scales, standard gamble and time trade-off techniques are reliable and valid methods for eliciting scores for health and non-health related preferences (coefficient of internal reliability  $r$ : rating scale  $r$  ranges from 0.86-0.94; standard gamble  $r$  ranges from 0.77-0.92; and, time trade-off  $r$  ranges from 0.77-0.88) (Torrance, 1986), a cognitive burden exists for respondents when using standard gamble and time trade-off techniques to elicit preferences. Both techniques require the need for short-term memory, concentration and focus (Feeny et al., 1996). Consequently, it is argued, the initial use of a rating scale or feeling thermometer will assist respondents in their introspection concerning preferences for the various states being evaluated. States to be evaluated may include states the respondent has experienced, is now experiencing or has never experienced (hypothetical states) Feeny et al. (1996) state:

“The utility approach, when it includes hypothetical states, allows the investigator to obtain important information from all patients on how they



think they would feel if they experienced some of the infrequent outcomes. Evidence to date shows that evaluations by persons experiencing the state and by others for whom the state is hypothetical usually do not differ substantially” (p.90).

Some argue (Levin & McEwan, 2001; Torrance, 1976) that since the standard gamble technique obeys the axioms of Von Neumann-Morgenstern utility theory for decisions under uncertainty, the validity of other techniques may be determined by comparing results with the standard gamble technique. By contrast, many scholars (particularly those involved in the “nuts and bolts” of conducting these analyses) do not make recommendations in favour of one method over the other (see Clemen, 1996; Levin & McEwan, 2001; Lipsey & Wilson, 2000; Von Winterfield & Edwards, 1986). Clemen (1996) and Von Winterfield and Edwards (1986) argue that all methods are potentially error-prone, where no method has the potential to provide a completely reliable indication of individual preferences. However, as stated by Levin and McEwan (2000): “all of the methods are helpful ways of eliciting information about individual preferences” (p.206). Therefore, we can safely argue that the choice of a particular method will be dependent upon not only the types of preferences we want to elicit but also time and budget allocations. Additionally, given the uncertainties surrounding the reliability, validity and precision of utility values, it is important to perform a sensitivity analyses regardless.

As stated in Section 7.4, in this thesis we adapt the analytical hierarchy process method (AHP) to identify relative utilities values with respect to non health-related outcomes resulting from early childhood intervention programs. The AHP

method elicits preferences from people on a relative scale from 0 to 9, whereby Saaty (1980) uses mathematical operations in treating a pair-wise comparison matrix as if they are cardinal numbers. However, as argued in Section 7.4, provided the vector of priorities and preferences are interpreted as relative preferences, the use of cardinal number operations to derive the vector of priorities is mathematically sound. Moreover, the AHP method provides a sound, and proven technique that can be used for assisting policy-makers in making more structured decisions based on program outcome/s at a given life phase (e.g. childhood, adolescence); and, a method for determining the relative utility of non health-related outcomes and their respective indicators. Consequently, we argue that the methodology the AHP technique developed by Saaty (1980) will assist in providing solutions to two dilemmas faced by policy-makers.

### **7.7: The analytical hierarchy process (AHP) and preferences**

The Analytical Hierarchy Process (AHP) is a widely used multi-attribute decision making tool, designed to aid in solving complex multiple criteria problems in a number of application domains (Saaty, 1990, 2000). The origins AHP can be traced back to the 1970's (Saaty, 1977) where it was applied to solve difficult decisions regarding scarce resource allocation and planning needs for the military (Saaty, 1980). Since the 1980's, it has been applied in multiple contexts which range from the analysis of conflicts (e.g. Johannessen et al., 2004; Tarbell & Saaty, 1980) in which attempts are made to identify and analyse potential political structures that may serve as a resolution to a conflict, to forming corporate strategies to bolster effective competition (Saaty, 1994), and the adoption of a model for analysing facility location selection decisions (Yang & Lee, 1997). The method has also been employed in the

health sector to assist in choosing the best hospice model to aid in caring for the terminally ill (Saaty, 1994), and has been extended to incorporate the analysis of future education in the United States of America (Alexander & Saaty, 1977a), and applications in finance (Saaty, 1990).

AHP is by no means an unproven method, it has been used extensively over the past thirty years to aid decision-makers analyse difficult problems that require multi-criteria analysis. Through trade-offs, AHP identifies relative preferences for various decision options/alternatives under circumstances of uncertainty. In short, the decision-maker judges the importance of each criterion in pair-wise comparisons (expressed by posing the questions “which of the indicators is more important?” and secondly, “by how much?”), which is expressed on a semantic scale of 1 (equality) to 9 (i.e. an indicator may be voted to be 9 times more important than the one to which it is being compared). From these pair-wise comparisons, the relative preference weights of the various indicators are then calculated using an eigenvector technique. This results in a prioritised ranking, weighting or preference value for each decision alternative (Saaty, 2000).

Alexander and Saaty (1977a; Alexander & Saaty, 1977b), in relation to a conflict in Northern Ireland, used the AHP methods as both a forward and backward process to move from one point to another to determine the likely outcome of a conflict, or to determine the necessary actions to achieve a desired state (Saaty, 1990).

Alexander and Saaty (1977a) first investigated the forward process to identify the likely political structures to emerge as a resolution to a conflict, given how the

outcomes from the alternative structures satisfy a set of parties to the conflict and their objectives. Johannessen, Bandara and Smith (2004) posit:

The forward process can be regarded as a one point boundary problem fixed at the present state. In practice, Alexander and Saaty were interested in determining the likely political structure that may serve as a resolution to a conflict, given the present state of a set of parties and their current objectives, capabilities and policies. Such a process is descriptive, and is referred to as the forward process of conflict analysis (p.165)

Alternatively, the backward process is a normative process that states a desired future outcome, and determines what actions are required to achieve that outcome (Alexander & Saaty, 1977b). The backward process can be considered as a one-point boundary problem fixed in the future. Thus, the objective is to remove obstacles and identify and introduce policies that would effectively produce the desired outcome/s (Saaty, 1990, 2000).

## **7.8: Conclusion**

In this chapter we have examined the value-of-life literature, revealing obstacles to valuing alternative life courses. Further, with the aim of drawing closer to developing a methodological standard that can value alternative life courses, we discussed the literature on cost-effectiveness and cost-utility analysis. Further, a detailed discussion has been provided regarding the concept of ‘utility’, specifically focussing on the individual utility function, social utility functions, and Pareto Optimality. Various methods of eliciting preferences have been explored including: rating scales and their variants; standard gamble (health related); time trade-off

(health related); and, multi-attribute utility. Additionally, a discussion regarding the best method to apply when evaluating non health-related outcomes, with the goal of providing a solution to the two dilemmas faced by policy makers was conducted. Consequently, we identified and briefly introduced the analytical hierarchy process method, which provides a structured method for analysing complex, multi-criteria problems, and eliciting relative utilities for outcomes associated with early childhood intervention programs on non health-related outcomes throughout an individual's life course.

In the following chapter (Chapter 8) a detailed discussion is provided of the analytic hierarchy process. This includes a discussion regarding organising complex structures into a systematic problem-solving procedure, the use of hierarchies in modelling decisions under uncertainty, decomposition (the structuring of a hierarchy to capture all the elements of a given problem), comparative judgements (the development of a matrix to perform pair-wise comparisons of the relative importance of the proposed elements) and the synthesis of priorities (calculating estimates of the vector, calculation of the consistency index and random index). Additionally, the benefits and criticisms of the AHP method are outlined. An illustrative example of Saaty's AHP is also provided.

## **CHAPTER 8: THE ANALYTIC HIERARCHY PROCESS (AHP): AN OVERVIEW**

### **8.1: Introduction**

In Chapter 1, it was argued that two methodological gaps exist with respect to policy decision-making and the economic evaluation of early childhood/crime prevention intervention programs. The first gap related to the array of tools accessible to make well informed choices on resource allocation and structured decision-making with respect to alternative policy options for early childhood interventions. The second related to our inability, using current methods, to measure the economic impact of early childhood interventions on outcomes associated with non-health related quality of life. Moreover, in Chapter 7, we argued that the AHP method provides an approach for use in closing both gaps in uncertain circumstances, and measuring preferences for salient outcomes across the life course (Saaty, 1980). Further, with these preference values we are able to conduct cost-utility analysis of competing and often disparate early childhood intervention programs ensuring that a set of common metric outcomes exist.

This chapter provides a detailed discussion of analytic hierarchy process. Section 8.2 focuses on organising complex structures into a systematic problem-solving procedure. It incorporates a discussion on the use of hierarchies in modelling decisions under uncertainty. Issues regarding decomposition (the structuring of a hierarchy to capture all the elements of a given problem), comparative judgements (the development of a matrix to perform pair-wise comparisons of the relative importance of the proposed elements) and the synthesis of priorities (calculating

estimates of the vector, calculation of the consistency index and random index) are introduced. The benefits and criticisms of the AHP method are outlined in Section 8.3. This section also includes a discussion of the axiomatic foundations of the AHP method. An illustrative example of Saaty's AHP is provided in Section 8.4. Finally, some conclusions are drawn in Section 8.5.

## **8.2: Organising complex decisions into a hierarchical framework**

Examining multiple dimensions of a problem at one time is a difficult task for most. Given that human perception and judgement are subject to change when psychological states or information inputs vary, Saaty (1980) proposed a new approach. He argued the need to develop a system, whereby, instead of organising our problems into complex structures, a conceptually simple approach could be adopted, which would nevertheless capture enough of the essence of the problem to assist in real world decisions and their complexities in relation to this problem. Thus, Saaty's analytic hierarchy approach provides a systematic procedure for representing the elements of a problem, rationally disaggregating the elements into smaller constituent parts, and introducing simple pair-wise comparison judgements for use in developing preference weights for priority ranking alternatives (Alexander & Saaty, 1977a).

Saaty (1986) argues that his AHP method involves three critical components: (1) decomposition, (2) comparative judgements, and (3) synthesis of priorities.

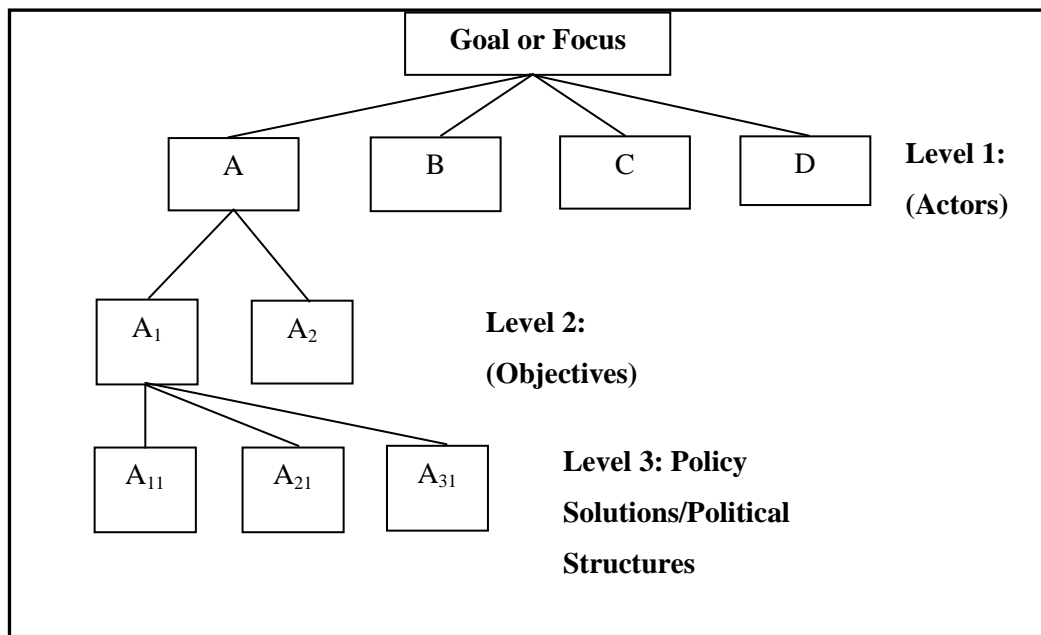
### 8.2.1: Decomposition of decision

Decomposition involves the structuring of a hierarchy to capture all the elements of a given problem. The use of hierarchies is a practical tool for modelling a conflict/problem (using either the forward or backward process of conflict analysis (as discussed in Chapter 7)) as it helps represent the elements of interest in relation to a given conflict problem. A hierarchical structure aids in representing the relationship among the elements at the different levels within the hierarchy (Isard & Smith, 1982; Johannessen et al., 2004). A hierarchy, as described by Saaty (1990) is comprised of a number of levels. For example, level 1 (the highest level) represents the key actors involved in the conflict/problem; the second level represents the key objectives of each of the actors; and level 3 provides the ‘political’ structures that may serve as future solutions to the conflict. Many more levels may be included in hierarchies representing any given policy dilemma or conflict situation. Figure 8.1 serves as a visual illustration of a hierarchical structure of a given problem. The figure demonstrates that parties A, B, C, D constitute level 1 of the hierarchy. Level 2, represents each party’s objectives ( $A_1, A_2 \dots A_n$ ) and level 3 is representative of the political structures ( $A_{21} \dots A_{n1}$ ).

### 8.2.2: Comparative judgements within the hierarchy

The comparative judgement component requires that one develop a matrix in order to perform pair-wise comparisons of the relative importance of the elements at each level of the hierarchy, developed for the particular problem being analysed. The scale proposed by Saaty for entering judgements or preference statements is presented in Table 8.1.



**Figure 8.1: An example of a hierarchy**

Source: Alexander, J. M., & Saaty, T. L. (1977a). The forward and backward process of conflict analysis. *Behavioural Science*, 22, 87-98.

**Table 8.1: Saaty's Comparison scale**

<i>Intensity of Importance</i>	<i>Definition</i>	<i>Explanation</i>
1	Equal importance	Two elements are of equal importance
3	Weak importance	Experience and judgement slightly favour one element over another
5	Essential or strong importance	Experience and judgement strongly favour one element over another
7	Demonstrated or very strong importance	An element is strongly favoured and its dominance is demonstrated in practice
9	Absolute importance	The evidence favouring one element over another is of the highest possible affirmation
<b>2,4,6,8</b>	Intermediate values	When compromise is needed

Adapted from: Saaty, T. L. (1990). *Multicriteria decision making: The Analytical Hierarchy Process*. Pittsburgh, USA: RWS Publications.

Pair-wise comparisons between elements are made on a scale from 1 to 9, where one indicates that two variables are equally important and nine indicates that only one variable is of absolute importance to another (see Table 8.1). For example, consider a pair-wise comparison matrix that relates to evaluating how

powerful/influential four parties are to resolution of a given conflict (parties A, B, C, D) (Table 8.2). Numbers in positions (A, A), (B, B), (C, C) and (D, D) are equal to one, implying that each of the parties is equally important to itself. The number five in position (A, B) implies that party A is considered to be strongly more important than party B in terms of influencing a successful outcome for the conflict problem. The value  $1/5$  in position (B, A) is simply the reciprocal of position (A, B). A key concern here is to ensure that the relative statements are consistent throughout the matrix given in Table 8.2 (Saaty, 1990, 2000).

**Table 8.2: Example of pairwise comparisons of parties (A, B, C and D)**

<b>Relative importance</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
<b>A</b>	<b>1</b>	<b>5</b>	<b>6</b>	<b>7</b>
<b>B</b>	<b><math>1/5</math></b>	<b>1</b>	<b>4</b>	<b>6</b>
<b>C</b>	<b><math>1/6</math></b>	<b><math>1/4</math></b>	<b>1</b>	<b>4</b>
<b>D</b>	<b><math>1/7</math></b>	<b><math>1/6</math></b>	<b><math>1/4</math></b>	<b>1</b>

Source: Saaty, T. L. (1990). *Multi-criteria decision making: The Analytical Hierarchy Process*. Pittsburgh, USA: RWS Publications.

Additional matrices are then developed involving pair-wise comparison of the elements of the second level (see Figure 8.1) with respect to the appropriate ‘parents’ in the first level. This process of developing appropriate pair-wise comparison matrices continues down the hierarchy (Saaty, 1990). As stated earlier, one could start from the top of the hierarchy and work down (forward process), or start from the bottom and work up (backward process). The third component involves the synthesis of the priorities inherent in these various pair-wise comparison matrices.

### 8.2.3: Synthesis of priorities

The synthesis of priorities occurs from level 1 down (see Figure 8.1); whereby “one multiplies local priorities by the priority of the corresponding criterion in the level above, and adding them for each element in a level according to the criteria it affects” (Saaty, 1986, p.842). Calculating estimates of the resulting vector of priorities can be obtained in the following four ways: (1) the crudest method, sums elements in each row and normalises by dividing each sum by the total of the sums; (2) a better method involves taking the sum of the elements in each column and form the reciprocals of these sums. To normalise, add to unity, divide each reciprocal by the sum of the reciprocals; (3) an improvement on the last method involves dividing the elements of each column by the sum of that column (i.e. normalising the column) and then add the elements in each resulting row and dividing the sum by the number of elements in each row; (4) a final method requires the multiplication of  $n$  elements in each row and take the  $n$ th root, and then normalisation of the resulting numbers (Saaty, 1990).

To illustrate further the differences between these four methods, we will apply them to the results inherent in the pair-wise comparison matrix given in Table 8.2. Applying method (1) would result in column vectors (19.00, 11.20, 5.42, 1.56). The total sum of the matrix (37.18) is calculated by summing the above vector components. If one divides each component of the vector by this number we obtain a column vector of priorities (0.51, 0.30, 0.15, 0.04) for options A, B, C, D respectively. Method (2) involves summing the columns of the matrix to obtain row vectors (1.51, 6.43, 11.25, 18.00). Reciprocals of these sums are (0.66, 0.16, 0.09, 0.06). When normalised they become (0.68, 0.16, 0.09, 0.06). Method 3 involves normalising each

column of the matrix (add its components and divide each component by the sum) to obtain the following matrix:

$$\begin{bmatrix} 0.66 & 0.78 & 0.53 & 0.39 \\ 0.13 & 0.16 & 0.36 & 0.33 \\ 0.11 & 0.04 & 0.09 & 0.22 \\ 0.09 & 0.03 & 0.02 & 0.06 \end{bmatrix} \rightarrow \begin{bmatrix} 2.36 \\ 0.98 \\ 0.46 \\ 0.20 \end{bmatrix} \xrightarrow{\text{Divided by 4}} \begin{bmatrix} 0.590 \\ 0.245 \\ 0.115 \\ 0.050 \end{bmatrix}$$

By summing the rows we obtain a column vector (2.36, 0.98, 0.46, 0.20), which when averaged by the sample size of four columns generates the column vector of priorities (0.590, 0.245, 0.115, 0.050) resulting from use of method (3). When multiplying  $n$  elements in each row, taking the  $n$ th root, and normalising the resulting numbers we obtain the results of method (4), namely (0.61, 0.24, 0.10, 0.05). In other words, to obtain the exact solution to a problem, we raise the matrix to arbitrarily large powers, and divide the sum of each row by the sum of the elements of the matrix (Saaty, 1990, 2000). Moving from method one to method four improves the result. If all entries in the initial matrix are consistent, the four methods should generate the same result. However, if entries are inconsistent, only method four offers an accurate approximation to the relative importance or priority assigned to each element. Method four is also known as the eigenvalue method. As argued by Saaty (1990), a normalised vector reveals the relative importance of or preference for the elements in the matrix. The eigenvalues combined with the use of hierarchies allow researchers to compare alternative solutions in relation to a conflict/problem.

Given that people are often inconsistent in answering questions, an important feature of AHP is to calculate the consistency level of the estimated vector of

priorities (or preference values). In particular, a consistency ratio (CR) can be calculated and used to measure the consistency in the pair-wise comparison. Saaty (1994) argues that one can identify acceptable CR values for different matrices' sizes as follows:

- 0.05 for a 3-by-3 matrix;
- 0.08 for a 4-by-4 matrix; and,
- 0.1 for larger matrices.

One should note that these values represent an upper limit; therefore, any values beyond these limits indicate inconsistent responses or intransitive ordering of preferences. A transitive ordering of preferences requires that if  $X$  is preferred to  $Y$ , and  $Y$  is preferred to  $Z$ , then  $X$  is preferred to  $Z$ . This logic of transitivity is taken as an axiom of rationality with regards to preferences of individuals (Boardman et al., 1996, 2001; Boardman, Greenberg, Vining, & Weimer, 2006). A violation of transitivity implies an ambiguous ordering of alternatives. Thus, if the CR value falls within the acceptable range then the resulting vector of priorities can be validly used to analyse the problem (Saaty, 1994).

To derive the consistency ratio (CR), we must firstly calculate  $\lambda_{\max}$ . To calculate  $\lambda_{\max}$ , we first multiply the matrix of pair-wise comparisons on the right by the estimated solution vector, to obtain a new vector.

Estimated solution  
vector

New vector

$$\begin{bmatrix} 1 & 5 & 6 & 7 \\ 1/5 & 1 & 4 & 6 \\ 1/6 & 1/4 & 1 & 4 \\ 1/7 & 1/6 & 1/4 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 0.59 \\ 0.25 \\ 0.11 \\ 0.05 \end{bmatrix}$$

$$\rightarrow$$

$$\begin{bmatrix} 1x.59 + & 5x.25 + & 6x.11 + & 7x.05 \\ .2x.59 + & 1x.25 + & 4x.11 + & 6x.05 \\ .17x.59 + & .25x.25 + & 1x.11 + & 4x.05 \\ .14x.59 + & .17x.25 + & .25x.11 + & 1x.05 \end{bmatrix}$$

$$=$$

$$\begin{bmatrix} 2.85 \\ 1.11 \\ 0.47 \\ 0.20 \end{bmatrix}$$

Divide the first component of this vector by the first component of the estimated solution vector, the second component of the new vector by the second component of the estimated solution vector and so on, to obtain another vector. Take the sum of components of this vector and divide by the number of components. The closer the solution ( $\lambda_{\max}$ ) is to  $n$  (number of activities in the matrix) the more consistent the result (Saaty, 1990).

$$\begin{bmatrix} 2.85 \\ 1.11 \\ 0.47 \\ 0.20 \end{bmatrix} \div \begin{bmatrix} 0.59 \\ 0.25 \\ 0.11 \\ 0.05 \end{bmatrix} = \begin{bmatrix} 4.83 \\ 4.44 \\ 4.27 \\ 4.00 \end{bmatrix} \rightarrow \lambda_{\max} = 17.54 \div 4 = 4.39$$

(17.54)

Next, we use our  $\lambda_{\max}$  value to derive a consistency index (C.I). C.I represented by  $(\lambda_{\max} - n)/(n-1)$ . In our numerical example,  $C.I = (4.39 - 4) / (4-1) = 0.13$ . The *random index* (R.I) provided below in Table 8.3, represents the consistency index (C.I) of a randomly generated reciprocal matrix with reciprocals forced.

**Table 8.3: Average random index (R.I) for matrices of order 1-15**

Order of matrix	1	2	3	4	5	6	7	8	9	10	11	13	14	15
Average R.I	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.56	1.57	1.59

The consistency ratio (CR) is equal to C.I/R.I, or in our numerical example,  $CR=0.13/0.90 = 0.14$ . This represents an unacceptably high CR for a 4 x 4 matrix. Consequently, it is argued that a method be chosen to reduce this value to a more consistent value, as proposed by Saaty (1994). Methods of improving consistency are discussed in detail in Chapter 9.

### 8.3: Benefits and criticisms of the AHP method

Saaty (2001) argues that the AHP method can be, and has been, applied to analyse real-world problems. It has been particularly useful for the allocation of resources (e.g. Cheng & Li, 2001a, 2001b), planning (e.g. Crowe, Noble, & Machimada, 1997; Udo, 2000; Yang & Lee, 1997), impact of policy (e.g. Saaty, 1980, 2001), and resolving conflicts (e.g. Johannessen et al., 2004; Tarbell & Saaty, 1980). The method is widely applied to inform corporate planning, portfolio selection, and cost-benefit analysis by government agencies for the purposes of resource allocation (Saaty, 2001). The advantages of the AHP method as an approach to problem solving and decision analysis are summarised in Figure 8.2. Benefits of the AHP method include: a flexible model for providing solutions to a range of unstructured problems; enabling people to refine their definition of a problem and to improve their judgement and understanding through repetition; integrating deductive and system approaches in solving complex problems; synthesising a representative outcome from diverse judgements; allowing interdependence of elements in a system, therefore not relying

on linear thinking; enabling people to select the best alternative according to their goals; allowing the natural tendency of the human mind to sort elements of a system into different levels and group like elements in each level; estimating, overall, the desirability of each alternative; providing a scale for measuring intangibles and a method for establishing priorities; and, tracking the logical consistency of judgements used in determining priorities.

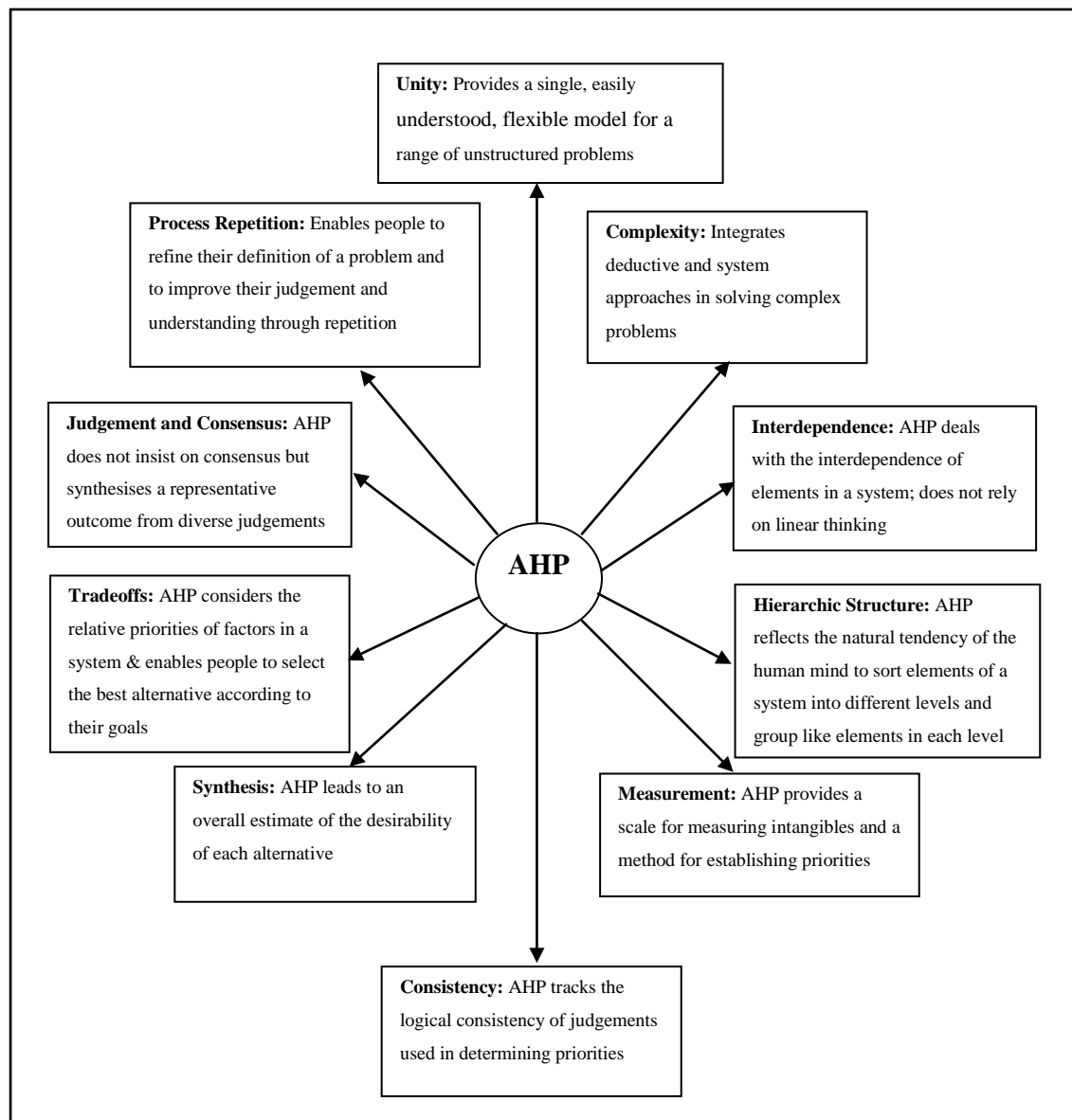
Since the inception of the AHP method (Saaty, 1977) a number of criticisms have been raised. Harker and Vargas (1987) in their analysis of the theoretical issues of the AHP method identified four main criticisms: the lack of an axiomatic foundation; ambiguity of the questions that the decision maker must answer; the scale used to measure the intensity of preference; and, the principle of hierarchical composition and rank reversal. Each of these criticisms is discussed in more detail below.

### **8.3.1: Criticism 1 - Axiomatic foundations of AHP**

In response to the first category of criticisms (the lack of an axiomatic foundation), Saaty (1986) elaborated the axiomatic basis of his approach. This is summarised in Appendix C by spelling out clearly the sound mathematical underpinnings of the method. Moreover, the implied assumptions relating to preference statements required from participants are highlighted. Consequently, this first area of criticism has now been laid to rest.



**Figure 8.2: Advantages of the analytical hierarchy process (AHP) method**



Source: Saaty, T. L. (2001). *Decision making for leaders* (3rd ed.). Pittsburg, PA: RWS Publications (p.25).

### **8.3.2: Criticism 2 - Ambiguity of the questions that the decision-maker must answer**

The second area of criticism relating to the AHP method is that there is a degree of ambiguity in regards to the question that the decision maker must answer when eliciting pair-wise comparisons. Watson and Freeling (1982) argue:

One present criticism of the AHP is concerned only with this stage of analysis [of applying the method to real-world problems]...What sort of question needs to be asked to elicit the numbers in this matrix [of pair-wise comparisons]? It would seem that they have to be of the form: “Which is more significant, purchase price or maintenance cost per year?”...If this question is asked without further explanation, it is, we maintain, meaningless (p. 583).

However, as argued by Harker and Vargas (1987), the problem of ambiguity is not a flaw of the AHP process itself, but occurs as a result of the fundamental question concerning the frame of reference in which a respondent is asked to make the necessary subjective judgements when implementing the AHP process to analyse a particular problem. Therefore, Harker and Vargas argue that the ambiguity of the question, or its lack of meaning, is dependent upon the cognitive environment in which one exists. They state:

One’s belief as to the meaning of terms such as “more important” or “strongly important” is a function of the cognitive frame of reference in which one currently resides. These definitions will vary from day to day and from individual to individual because this reference frame varies over time and individuals. While it is true that a poorly worded question yields poor results

and that better wording of a question can significantly increase the effectiveness of the methods, no method or no perfect question will ever remove ambiguity completely due to the reliance on the individual's frame of reference (p. 1387).

Moreover, criticisms raised by Wason and Freeling (1983) lacked a comprehension of the 'frame of reference' or an understanding that ambiguity not explicable within the context of the frame of reference is not a failure of the method, but a failure of the analyst or decision-maker to understand the issue at hand and state questions which are meaningful in the context of the frame of reference of respondents (Harker & Vargas, 1987). Overall, Harker and Vargas (1987) argue that ambiguity is a phenomenon that can be present in all preference eliciting methods. Therefore, clear and concise definitions of the criteria, subcriteria and alternatives are essential when eliciting preferences within any particular application of the AHP method. For a more detailed discussion of this argument, see Harker and Vargas (1987).

### **8.3.3: Criticism 3 - The scale used to measure the intensity of preference**

Another controversial aspect of the AHP method is the scale used to measure the intensity of preferences  $P_c(A_i, A_j)$ , between the alternatives  $A_i, A_j \in A$ . Manipulation of this ratio scale within the method implicitly assumes that the cardinal intensity of preferences between the alternatives can be expressed by using the scale. Critics of the method (e.g. Dyer, 1990; Dyer & Wendell, 1985) have questioned the ability of Saaty's 1 to 9 scale to accurately measure this cardinal intensity of preferences. They in effect argue that respondents believe they are giving relative

preference statements (a ratio scale), yet these are subsequently manipulated mathematically in a way which appears to be treating them as cardinal preference statements. However, Harker and Vargas (1990) have demonstrated that although the scale of measure may be a subject of debate, the mathematics behind the AHP follow directly from the axiomatic foundations. Harker and Vargas (1987) provide a detailed overview of the debate regarding the use of Saaty's 1 to 9 scale to measure the intensity of preferences. They argue that, in practice, the 1 to 9 scale can accurately portray an individual's intensity of preference. They state:

The 1 to 9 scale may be able to capture a great deal of information and has proven to be extremely useful due to the fact that the AHP is somewhat scale independent. The AHP, through the use of the eigenvector and normalization procedures, is a highly nonlinear operator on the scale of measure. Thus, a simple linear 1 to 9 scale can easily represent a highly nonlinear cognitive scale when the AHP is used in conjunction with this linear scale. Furthermore, the AHP requires an individual to make  $n(n - 1)/2$  pair-wise comparisons when comparing  $n$  alternatives, versus the  $(n - 1)$  comparisons which would have been made if cardinal transitivity (or consistency) is enforced. The judgements obtained by using the AHP contain redundant information so that even if any one judgements forced to be inconsistent due to the 1 to 9 scale, the final weights are not substantially affected (pp.1387-1388).

This argument is supported by research conducted by Vargas (1984) on the stability of positive reciprocal matrices, which demonstrated that any inconsistencies caused by the 1 to 9 scale do not significantly affect the final weights.

Harker and Vargas argue that in theory, any ratio scale could be used in the AHP method; however, it is recommended that until a consensus is reached by the research community regarding the scale numerical values via experimentation with various scales of measure, Saaty's "proven" 1 to 9 scale should be adopted.

#### **8.3.4: Criticism 4 - The principle of hierarchical composition and rank reversal**

Belton and Gear (1984) argue that the notion of hierarchical composition of levels of criteria and composition are firstly, too simplistic, and secondly that rank reversals could occur. Rank reversal is a phenomenon whereby there is the possibility of changes in the relative rankings of the other alternatives after an alternative is included or excluded (Wang & Elhag, 2006).

With respect to the first criticism, Harker and Vargas (1987) argue that Belton and Gear's postulation that the notion of hierarchical composition of levels of criteria and composition are too simplistic is ill-founded. Harker and Vargas (1987) provide a simple example to support their claims for legitimacy of the technique. Consider a hierarchy composed of three criteria  $C_1$ ,  $C_2$  and  $C_3$  with two alternatives  $A_1$ ,  $A_2$  which are compared with respect to  $C_1$ ,  $C_2$  and  $C_3$ . An assumption is made whereby the alternatives influence the criteria and the criteria also influence the alternatives. This is known as a system with feedback (Saaty, 1980). Consequently, the system is no longer a simple hierarchy but a complex network of interactions. For a detailed example of Saaty's system of feedback, see (Saaty, 1980) and (Harker & Vargas, 1987).

The second problem proposed by Belton and Gear (1984) concerning the validity of the AHP method involves rank reversal. The proposition, however, has

been demonstrated to be “vacuous” (Harker & Vargas, 1987, p.1400) as demonstrated in Appendix D.

In summary, Dyer (1990) argues that all criticisms raised regarding the AHP model have been proven to be operational in nature and not a flaw in the basic methodology. Further, Harker and Vargas (1987) highlight that some of the critiques have been due to, in part, a misunderstanding of the theoretical basis of the AHP and not serious methodological flaws.

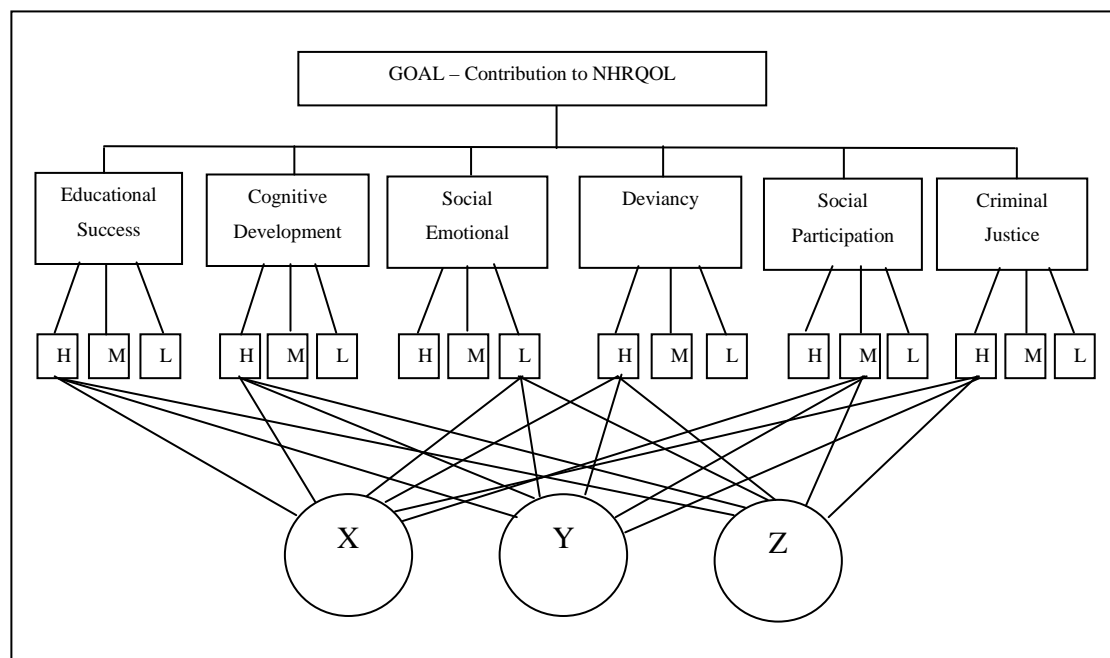
#### **8.4: An illustrative example of Saaty’s method**

We illustrate a working example of the AHP method on a policy decision model, whereby the utilisation of Saaty’s ‘backward process’ is adopted. As stated earlier, the AHP process is a simple and efficient technique for the analysis of decision problems. This section, therefore, provides a step-by-step procedure for attaining preference values for decisions regarding alternative early childhood intervention program options and their contribution to increasing non-health related quality of life during the adolescent life phase (goal of the decision process).

Consider a government who wants to determine preferences for three different forms of early childhood intervention. The attributes considered most relevant (based on objective research) from the decision-makers/policy-makers perspective are (1) educational success, (2) cognitive development, (3) social-emotional development, (4) deviancy, (5) social participation, and (6) criminal justice outcomes. Three forms of early childhood intervention (Program X, Program Y and Program Z) possess all these attributes, but with the possibility of different levels of intensity (outcome levels of

success): high (H), medium (M), and low (L). Based on the concept of “bounded rationality”, that is, “consumers [individuals] do not act on perfect or complete information and are satisfied with less than the economically most rational choice” (Saaty, 2001, p.102) – attributes may be divided into a smaller number of intensity categories to help distinguish between the large number of attributes (Saaty, 2001). Consequently, the hierarchy depicted in Figure 8.3 can be established.

**Figure 8.3: An illustrative hierarchy for determining preferences**



The problem of selecting the early intervention program with the greatest overall preference can be analysed in the following manner:

**Step 1:** First, we develop a matrix that compares attributes in pairs with respect to the goal (most preferred early intervention program) with the aim of determining preferences among the attributes (see Figure 8.4). Individual responses are aggregated and an unweighted average calculated (Individual weightings will result in a slightly different result). Additionally, it is argued whether a weighted

scheme is appropriate, particularly when respondents are seen as having equally important contributions to the decision-making process (in this case experts). A discussion seen as beyond the scope of this thesis. We then normalise each column of the matrix by dividing each preference score by the total of its column (see Figure 8.5.) For example, in our numerical example we divide 1 (representing the preference score of ES by ES) by the total of column 1 (20.20) to obtain a normalised entry of 0.05. We then add the elements of each row to obtain a new column vector (see Figure 8.5) To attain a column vector of priorities for the attributes, we then divide each entry in the sum of normalised components column vector by the number of attributes, namely 6 (see Figure 8.6). A detailed procedure for performing this function is outlined in Section 8.2.

**Figure 8.4: Matrix comparing six outcome domains**

Goal	ES	CD	SED	D	SP	CJO
ES	1	1/4	1/5	1/4	5	1/6
CD	4	1	1/3	3	6	1/2
SED	5	3	1	4	7	3
D	4	1/3	1/4	1	5	1/5
SP	1/5	1/6	1/7	1/5	1	1/7
CJO	6	2	1/3	5	7	1
	20.20	6.75	2.25	13.45	31	5.01

New Row Vector

←



**Figure 8.5: Normalised columns of the matrix**

Goal	ES	CD	SED	D	SP	CJO	
ES	0.05	0.04	0.09	0.02	0.16	0.03	<div style="display: flex; align-items: center;"> <div style="border: 1px solid black; padding: 2px; margin-right: 10px;">New Column Vector</div> <div style="display: flex; flex-direction: column; align-items: center;"> <div style="margin-bottom: 10px;">↙</div> <div style="margin-bottom: 10px;">→</div> </div> <div style="display: flex; flex-direction: column; align-items: center;"> <div style="margin-bottom: 10px;"><math>\begin{bmatrix} 0.39 \\ 1.01 \\ 2.26 \\ 0.63 \\ 0.17 \\ 1.55 \end{bmatrix}</math></div> </div> </div>
CD	0.20	0.15	0.15	0.22	0.19	0.10	
SED	0.25	0.44	0.44	0.30	0.23	0.60	
D	0.20	0.05	0.11	0.07	0.16	0.04	
SP	0.01	0.03	0.06	0.01	0.03	0.03	
CJO	0.30	0.30	0.15	0.37	0.23	0.20	

**Figure 8.6: Obtaining the vector of priorities**

$$\begin{bmatrix} 0.39 \\ 1.01 \\ 2.26 \\ 0.63 \\ 0.17 \\ 1.55 \end{bmatrix} \div 6 \longrightarrow \begin{bmatrix} 0.06 \\ 0.17 \\ 0.38 \\ 0.10 \\ 0.03 \\ 0.26 \end{bmatrix}$$

Vector of priorities for Level 1  
of the hierarchy

This new vector highlights that, with a value of 0.38, the attribute SED is the most preferred attribute with respect to its potential to contribute to our goal (contribution to non health-related quality of life). This is followed in perceived importance by the attributes CJO (0.26), CD (0.17), D (0.10), ES (0.06) and SP (0.03). Moreover, each attribute now has a perceived preference value which may be used in cost-utility applications (see Chapter 7 and Section 8.1).

**Step 2:** To determine preferences among the intensities of the attributes, we develop six matrices that compare the intensity levels (outcome levels of success) (high, medium and low) in pairs with respect to all six attributes (educational success,

cognitive development, social-emotional development, deviancy, social participation and criminal justice outcomes) (see Figures 8.7, 8.8, 8.9, 8.10, 8.11 and 8.12).

**Figure 8.7: Determining preferences among the intensities for the outcome educational success (ES)**

ES	H	M	L		
H	1	5	8		
M	1/5	1	5	→	0.73
L	1/8	1/5	1		0.21
					0.06

← Priority vector

$$\lambda_{\max} = 3.15, \text{C.I} = 0.07$$

**Figure 8.8: Determining preferences among the intensities for the outcome cognitive development (CD)**

CD	H	M	L	
H	1	7	9	
M	1/7	1	7	→
L	1/9	1/7	1	

Priority

0.76

0.19

0.05

$$\lambda_{\max} = 3.33, \text{C.I} = 0.16$$

**Figure 8.9: Determining preferences among the intensities for the outcome social-emotional development (SED)**

SED	H	M	L	
H	1	1/7	1/9	
M	7	1	1/7	→
L	9	7	1	

Priority

0.05

0.19

0.76

$$\lambda_{\max} = 3.33, \text{C.I} = 0.16$$

**Figure 8.10: Determining preferences among the intensities for the outcome deviancy (D)**

D	H	M	L		Priority
H	1	3	5	→	0.63
M	1/3	1	4		0.28
L	1/5	1/4	1		0.09

$$\lambda_{\max} = 3.09, \text{C.I} = 0.04$$

**Figure 8.11: Determining preferences among the intensities for the outcome social participation (SP)**

SP	H	M	L		Priority	
H	1	1/5	2	→	0.18	$\lambda_{\max} = 3.05, \text{C.I} = 0.03$
M	5	1	5		0.71	
L	1/2	1/5	1		0.11	

**Figure 8.12: Determining preferences among the intensities for the outcome criminal justice outcomes (CJO)**

CJO	H	M	L		Priority
H	1	7	9	→	0.76
M	1/7	1	7		0.19
L	1/9	1/7	1		0.05

$$\lambda_{\max} = 3.33, \text{C.I} = 0.16$$

We then develop a matrix combining the vectors of priority or preferences for intensities of each attribute. For example, the priority vectors for Figures 8.7 – 8.12 are combined into one matrix (see Figure 8.13). Across the top of this matrix we have listed the priority values derived from Figure 8.6 highlighting priority values for each attribute (e.g. ES, CD, SED, D, SP, CJO).

**Figure 8.13: Matrix comparing intensity levels with respect to each outcome**

	(0.06)	(0.17)	(0.38)	(0.10)	(0.03)	(0.26)
	<i>ES</i>	<i>CD</i>	<i>SED</i>	<i>D</i>	<i>SP</i>	<i>CJO</i>
<i>H</i>	0.73	0.76	0.05	0.63	0.18	0.76
<i>M</i>	0.21	0.19	0.19	0.28	0.71	0.19
<i>L</i>	0.06	0.05	0.76	0.09	0.11	0.05

**Step 3:** In step 3, we multiply each column from Figure 8.13 by the priority of the corresponding attribute to obtain the weighted vectors of priority for the various outcome intensities (See Figure 8.14). For example, row 1/column 1 (H-ES)(0.73) x the priority of the attribute ES (Figure 8.6) (0.06) = 0.041.

**Figure 8.14: Vectors of priority for the six outcome intensities**

	ES	CD	SED	D	SP	CJO
H	0.041	0.128	0.018	0.063	0.005	0.201
M	0.012	0.032	0.073	0.028	0.019	0.051
L	0.004	0.008	0.292	0.009	0.003	0.013

**Step 4:** Next, we select from each column the element with the highest priority to obtain the vector of desired attribute intensity/outcome (see Figure 8.15). It is possible that we do not need to select the element with the highest priority; however, failure to do so would make the next step of the method very complex and cumbersome. In other words, we would be faced with asking respondent to complete eighteen matrices of pair-wise comparisons compared to six. Saaty (2001) has used this method widely in his research and argues that it does not significantly impact on

the outcome. In fact, based on the concept of “bounded rationality”, individuals are able to better distinguish between large numbers of attributes when they are divided into a smaller number of intensity categories. This aids in achieving a more consistent group of answers.

**Figure 8.15: Normalised vector of priorities with respect to the most desired attribute-intensity outcome**

$$\begin{bmatrix} \text{H-ES} & \text{H-CD} & \text{L-SED} & \text{H-D} & \text{M-SP} & \text{H-CJO} \\ 0.041 & 0.128 & 0.292 & 0.063 & 0.019 & 0.201 \end{bmatrix}$$

Figure 8.15 demonstrates that a high effect is considered the most important level of intensity on the attributes ES, CD, D, and CJO. Further, a low effect is considered the most important level of intensity on the attribute SED, while a medium effect is considered to be the most important level of intensity for the attribute SP.

We then normalise this vector by dividing each component by the sum of the vector (see Figure 8.16). Normalising the vector will give us a vector of priorities with respect to the most desired attribute-intensity outcome.

**Figure 8.16: Normalising the vector**

$$\begin{bmatrix} 0.041 \\ 0.128 \\ 0.292 \\ 0.065 \\ 0.019 \\ 0.201 \\ \hline 0.746 \end{bmatrix} \xrightarrow{\text{Normalise}} \begin{bmatrix} 0.056 \\ 0.172 \\ 0.393 \\ 0.084 \\ 0.025 \\ 0.270 \end{bmatrix}$$

**Step 5:** Next we determine the perceived early childhood intervention options by developing matrices that compare the three options (X, Y and Z) in pairs with respect to the most desired attribute intensity outcomes (see Figures 8.17 to 8.22).

**Figure 8.17: Comparing options X, Y, and Z for desired attribute intensity H-ES**

H – ES	X	Y	Z		Priority	
X	1	5	8	→	0.715	← Priority vector
Y	1/5	1	5		0.219	
Z	1/8	1/5	1		0.067	

$$\lambda_{\max} = 3.18, \text{C.I} = 0.09$$

**Figure 8.18: Comparing options X, Y, and Z for desired attribute intensity H-CD**

H – CD	X	Y	Z		Priority
X	1	2	7	→	0.566
Y	1/2	1	8		0.373
Z	1/7	1/8	1		0.061

$$\lambda_{\max} = 3.08, \text{C.I} = 0.04$$

**Figure 8.19: Comparing options X, Y, and Z for desired attribute intensity L-SED**

L – SED	X	Y	Z		Priority
X	1	1/4	1/7	→	0.073
Y	4	1	1/5		0.205
Z	7	5	1		0.722

$$\lambda_{\max} = 3.12, \text{C.I} = 0.06$$

**Figure 8.20: Comparing options X, Y, and Z for desired attribute intensity H-D**

H – D	X	Y	Z		Priority
X	1	2	1	→	0.413
Y	1/2	1	1		0.260
Z	1	1	1		0.328

$$\lambda_{\max} = 3.05, \text{C.I} = 0.043$$

**Figure 8.21: Comparing options X, Y, and Z for desired attribute intensity M-SP**

M – SP	X	Y	Z		Priority
X	1	2	1	→	0.407
Y	1/2	1	3		0.369
Z	1	1/3	1		0.224

$$\lambda_{\max} = 3.37, \text{C.I} = 0.18$$

**Figure 8.22: Comparing options X, Y, and Z for desired attribute intensity H-CJO**

H – CJO	X	Y	Z		Priority
X	1	4	6	→	0.682
Y	1/4	1	4		0.236
Z	1/6	1/4	1		0.082

$$\lambda_{\max} = 3.11, \text{C.I} = 0.05$$

We then develop a matrix (Figure 8.23) combining the vectors of priority values for options X, Y and Z with respect to each attribute intensity. For example, the priority vectors for Figures 8.17 – 8.22 are combined into one matrix (see Figure 8.23). Across the top of this matrix we have listed the priority values derived from Figure 8.16 highlighting priority values for each attribute intensity/outcome (e.g. H-ES, H-CD, L-SED, H-D, M-SP, H-CJO).

**Figure 8.23: Overall option attributes perception - Matrix comparing attribute-intensity levels with respect to each alternative (X, Y, Z)**

	(0.056)	(0.172)	(0.393)	(0.084)	(0.025)	(0.270)
	H – ES	H – CD	L – SED	H – D	M – SP	H – CJO
X	0.715	0.566	0.073	0.413	0.407	0.682
Y	0.219	0.373	0.205	0.260	0.369	0.236
Z	0.067	0.061	0.722	0.328	0.224	0.082

We then multiply each column from Figure 8.23 by the priority of the corresponding attribute/intensity to obtain the weighted vectors of priority for the various option attribute perception (See Figure 8.24).

**Figure 8.24: Weighted overall option attribute perception**

	H – ES	H – CD	L – SED	H – D	M – SP	H – SR
X	0.039	0.097	0.029	0.035	0.010	0.184
Y	0.012	0.064	0.080	0.022	0.009	0.064
Z	0.004	0.010	0.283	0.028	0.006	0.022

*Step 7:* In our final step, we add each row from Figure 8.24 to obtain the overall priorities of the three alternatives. This synthesis yields the vector of priorities given in Figure 8.25

**Figure 8.25: Vector of priorities for program alternatives based on pair-wise comparisons of the all levels of the hierarchy**

$$\begin{bmatrix} \text{X} & 0.395 \\ \text{Y} & 0.252 \\ \text{Z} & 0.353 \end{bmatrix}$$

We conclude from Figure 8.25 that option X dominates options Y and Z on all desired attribute intensities (e.g. H-ES, H-CD). This information, therefore, provides a ranking of alternatives for policy-makers to make more informed complex multi-criteria decisions based on pair-wise comparisons of attributes (e.g. ES, CD, SED), outcome levels of intensity (e.g. H-ES, H-CD) and program options (X, Y and Z).



## 8.5: Conclusion

In this chapter we provided a detailed discussion of analytic hierarchy process (AHP). Issues regarding the organisation of complex structures into a systematic problem-solving procedure have been examined. Further, the benefits and criticisms of the AHP method have been outlined, including a discussion of the axiomatic foundations of the AHP method. Finally, an illustrative example of Saaty's AHP was provided.

The following chapter, Chapter 9, outlines the methodology applied in this thesis to first measure preference values, and secondly to solve complex multi-criteria problems using the AHP method. We begin the chapter by examining the need to make choices when information regarding longitudinal outcomes of early childhood interventions already exists. The remainder of the chapter is dedicated to outlining, in a series of steps, the adopted AHP method to first measure preference values for non health-related outcomes during adolescence and secondly, describe a structured method for making better policy decisions with respect to policy options for early childhood intervention program alternatives.

## **CHAPTER 9: DEVELOPING THE ANALYTIC HIERARCHY PROCESS FOR USE IN EVALUATION OF ALTERNATIVE EARLY-IN- LIFE INTERVENTION PROGRAMS**

### **9.1: Introduction**

This chapter outlines the methodology used to attain preference values for decisions regarding options for early childhood intervention programs and their contribution to increasing non-health related quality of life during the adolescent life phase. A methodology is provided for solving complex multiple criteria problems with the aim of providing a tool to aid in the policy planning process. Section 9.2 examines why choices are necessary when longitudinal research provides us with detailed information regarding the effectiveness of early childhood interventions on at-risk children and their families during the adolescent life phase. Section 9.3 outlines the method adopted in this thesis for determining preferences values with respect to non health-related outcomes associated with early childhood interventions for at-risk families and their children during the adolescent life phase; and, describes the particular analytical hierarchy process employed to achieve this outcome. Finally, some final remarks are provided in Section 9.4.

### **9.2: Why should we have to make a choice?**

Policy decisions regarding early childhood intervention/developmental crime prevention policies have traditionally relied on an unstructured mechanism for making decisions. Although policy decisions are often made with consideration of empirical research, we argue that complex decisions of this nature cannot be made adequately without a structured methodological process. Thus, the aim of this thesis is to provide

a systematic process enabling decisions to be made that incorporate all relevant criteria with respect to potential adolescent outcomes resulting from early childhood intervention programs. The adoption of this method allows us to make better and more informed choices regarding the development of policies to improve the lives of at-risk children and their families. One must acknowledge that decisions of this nature are extremely complex and cannot be effectively made without incorporating all relevant criteria. This is beyond our cognitive abilities; therefore, we argue that decisions made without a structured process are imperfect decision-making methods.

However, one can also question the need to make a choice regarding which early childhood intervention is the most preferred option with respect to its contribution to non health-related quality of life when longitudinal research highlights that combinations of programs produce better outcomes for those individuals most at-risk. This may be true; however, program funding is highly competitive and policy-makers are often left with the dilemma of which program to fund and how much funding a program or group of programs should receive. Thus, a method of making decisions that incorporates all relevant criteria with respect to this dilemma should be applied. In the event that a multi-component intervention is funded, it is argued that a structured approach is necessary to determine the disaggregation of funding that should be adopted.

Many variations of a hierarchy can be developed to analyse our particular policy dilemma. The development of this hierarchy depends on the question you are asking. For example, one may wish to know which outcome indicator is the most preferred option with respect to its contribution to the overall outcome domain. The

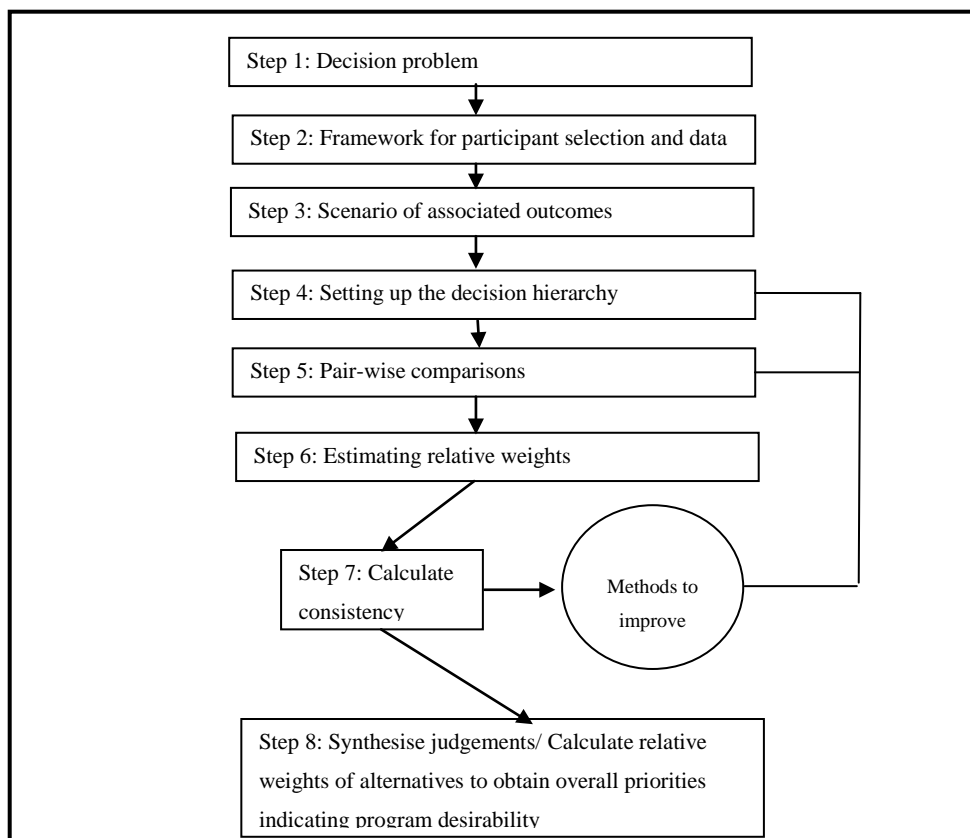
point is, hierarchical structures are only limited by ones imagination, and the question one is trying to answer (or goal you are seeking to achieve). As has been previously discussed, the question we are attempting to answer is which program alternative is most preferred with respect to its impact on non-health related quality of life during the adolescent life phase.

### **9.3: Application of the Analytical Hierarchy Process (AHP) Method**

Application of the AHP requires that one follow eight steps (see Figure 9.1). In step 1, the decision problem is clearly defined, together with identification of why the AHP method has been used to analyse this problem. In step 2 the framework to be employed for selection of respondents is developed. This includes targeting the participant groups that influence decisions regarding the implementation of program options available for early childhood intervention, methods employed for data collection from individual groups, and procedures adopted with respect to participant confidentiality. Step 3 puts together the information that participants are presented with prior to conducting the survey to attain preference values from respondents. This includes providing participants with evidence from empirical research to assist them in making more informed choices regarding preference values for the upcoming pair-wise comparisons. Step 4 involves the setting up of the decision hierarchy deemed most useful for analysing the given decision problem. Step 5 establishes the method used to elicit pair-wise comparisons, which allows participants to make judgements on the relative importance between two elements across all levels and associated elements of the hierarchy. Step 6 employs the method described in Section 8.4 for estimating the relative weights from pair-wise comparisons. This is followed in step 7 by an examination of participant responses and the adoption of a series of possible

methods for improving consistency. Finally, step 8 involves the synthesis of judgements and the calculation of relative weights of alternatives to obtain overall priorities indicating program desirability. Each of these steps is now discussed in more detail below.

**Figure 9.1: Steps in the application of the analytical hierarchy process (AHP) method**



Adapted from: Cheng, E. W. L., & Li, H. (2001). Analytical Hierarchy Process: An approach to determine measures for business performance. *Measuring Business Excellence*, 5(3), 30-36.

### 9.3.1: Step 1 - Decision problem

The existing developmental crime prevention literature does not provide preference values for early childhood intervention program options (e.g. structured preschool program, centre-based childcare/developmental day care, home visitation, family support services, and parental education), outcomes relating to these programs

during a given life phase (e.g. adolescence – educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes and family well-being) and their indicators (e.g. indicators of educational success include: special education services; feeling of belonging at school; school graduation; school drop-out; long-term school suspension; grade retention; and, completed years of education) or their various levels of success (e.g. no program impact, small impact, medium impact, high impact and very high impact). Such preference values could be used to identify the cost-utility of one program option compared to another.

It is the potential that the AHP method has to generate such preference values that provides the basis for the exploration of its usefulness in the evaluation of alternative early childhood intervention programs. For example, how do we choose among the alternatives available, the most preferred early childhood intervention program that has the most potential to reduce rates of delinquency during the adolescent life phase? Or, how do we choose among the alternatives available, the most preferred early childhood intervention program that has the most potential to increase family wellbeing during the adolescent life phase?

Application of AHP method to this problem relies on subjective judgements made by experts in the given field (e.g. policy makers, clinicians, academics etc). However, research demonstrates (Cheng & Li, 2001a, 2001b; Harker & Vargas, 1987; Saaty, 1980) that a simple rating method for eliciting preference statements from such experts cannot filter out the inconsistencies of responses (Cheng & Li, 2001a). In other words, a simple rating method cannot prevent respondents from providing their

answers arbitrarily, mistakenly, or carelessly. By contrast, as stated by Cheng and Li (2001a), “AHP is a structured method that can elicit biased options of decision makers in weighting and prioritization” (p.32). AHP achieves this by adopting a pair-wise comparison approach, where respondents are asked to compare two objects (e.g. educational success and cognitive development) at one time to formulate a judgement as to their relative weight (Saaty, 1990). This method has been proven (Harker & Vargas, 1987; Saaty, 1977, 2000) to be more accurate (with less experimental error) in achieving a higher level of consistency given that it requires respondents to think precisely before providing their answers. Saaty (1980) defines the term ‘inconsistency’ as a lack of transitivity of preferences. To test this, the AHP method employs a consistency test that can identify and screen out inconsistent responses. Moreover, the more a respondent is familiar with the situation, the more consistent the results will be (Cheng & Li, 2001a, 2001b).

The individuals who make up the study and control groups of early childhood intervention programs described in longitudinal research are considered those most at-risk. Individuals described in the longitudinal literature commonly face four prominent risk factors (1) community risks; (2) family risks; (3) peer risks; and (4) individual risks. The four risk factors are further disaggregated according to the literature (Durlak, 1998; Marshall & Watt, 1999; Pollard, Hawkins, & Arthur, 1999). The broad domain of community risks is disaggregated into individuals living within an impoverished neighbourhood, a neighbourhood that has higher than average crime rates, large rates of public housing, a neighbourhood with limited social service facilities, and a community that is culturally and linguistically diverse. Characteristics of family risks include: families with high levels of social and economic disadvantage

(e.g. lower than average median weekly household income, low rates of high school completion by parents, higher than average rates of single parent families and marital discord, punitive child-rearing practices, high unemployment, and socially isolated due to cultural and linguistic barriers). Peer risks are disaggregated into negative peer pressure or modelling and possible peer rejection. Finally, individual risks include early learning difficulties and possible early behavioural problems.

Five forms of early childhood intervention programs are analysed in this study. They include:

(a) *Structured preschool program*: A structured preschool program (SPP), typically centre-based, provides structured types of learning experiences within the context of developmentally orientated programs in which children work together with the preschool teacher to develop their language and cognitive skills. A structured preschool curriculum tends to incorporate mathematics concepts, language arts, reading readiness, social studies and science into structured classroom activities (Coalition for evidence-based policy, 2005; Parks, 2000; Weikart & Schweinhart, 1992).

(b) *Home visitation program*: Home visitation programs were developed to provide a variety of services to parents to reduce incidences of child abuse and neglect resulting from parents who lack the skills to properly care for an infant or without the social support networks required to support them as new parents (e.g. relatives or neighbours). Home visitation programs involve either a nurse, social worker, or neighbour visiting the homes of expectant mothers/parents to help motivate them as parents to learn and accept the help of others so that they might



provide a loving and nurturing environment for their new child (Olds et al., 1994).

Benefits of the program include:

- linking parents and infants to preventive medical care;
- providing resources to parents who are eager to learn the skills necessary to care for their child;
- promoting parent-child attachment and bonding;
- aiding parents in developing appropriate expectations for their child's development and providing resources to foster that development;
- providing support to families who have other younger children;
- identifying and providing guidance and support to already overburdened families; and
- facilitating the formation of long-term and trusting relationships between families and their support networks (James, 1995; Olds et al., 1997).

*(c) Centre-based childcare/Developmental day care:* Centre-based/developmental day care (CBCC) was developed to meet the growing demands for day care (due to the increasing numbers mothers returning to the workforce) coupled with the need to provide a developmental preschool program for children that would ultimately increase school readiness and assist children in attaining the social skills to function within society. The aim of developmental day care is to provide a stimulating and nurturing childcare environment for infants and young children between the ages of 6 weeks to 5 years of age (Reynolds, 1994). The goal of the majority of developmental day care centres is to:

- provide children with elementary academic readiness to promote optimal cognitive development and the social skills to successfully negotiate the early primary school experience;
  - provide the opportunity for children to develop gross motor skills;
  - instil a sense of values which will allow children to become responsible and constructive citizens;
  - provide parents with appropriate parenting skills and techniques;
  - act as a medium between families and various social service agencies;
- and,
- provide children with numerous cultural and real-life experiences
- (O'Brien Caughy et al., 1994; Reynolds, 1994; Reynolds et al., 2001).

*(d) Family/parenting support:* Family support services (FSS) work in partnership with parents to aid in crisis situations to minimise 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. Family support services offer one or more of the following: individual and family counselling, financial management, anger management, domestic abuse, sexual abuse, learning disability, and youth support services (Children's Home Society and Family Services, 2007).

*(e) Parent education program:* The primary aim of a parent education program is to improve core parenting skills. Sanders (2003) argues that aspects of core parenting skills include:

- improving parents' observation skills – e.g. monitoring both children's and one's own behaviour;
- developing parent-child relationship enhancement skills – e.g. spending quality time, showing affection, and talking with your child;
- encouraging desirable behaviour – e.g. providing activities which are engaging, developing the skill of giving non-verbal engagement, and providing descriptive praise;
- teaching new skills and behaviours – e.g. setting a good example and setting new developmentally appropriate goals;
- managing misbehaviour – e.g. establishing rules, adopting directed discussion; and giving clear, concise and calm instructions;
- preventing problems in high-risk situations – e.g. planning and preparation skills, discussing ground rules with the child, providing incentives and discussing consequences of inappropriate behaviour;
- self-regulation skills – e.g. self-evaluation of one's strengths and weaknesses, and setting personal goals;
- mood management and coping skills – e.g. relaxation and stress management skills, developing coping plans for difficult situations, and challenging unhelpful thoughts; and
- partner support and communication skills – e.g. improving personal communication, having casual conversations, supporting one another during difficult situations, problem solving, and improving relationships.

Seven outcomes relating to the above described programs during the adolescent life phase include: educational success, cognitive development, social-

emotional development, deviancy, social participation, criminal justice outcomes and family well-being. The outcomes described are all outcomes for at-risk/disadvantaged populations that have been described in the longitudinal research and presented in the meta-analysis in Chapter 6. Indicators relating to the seven outcome domains are given in Table 9.1.

**Table 9.1: Indicators of the outcomes educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes, and family wellbeing**

<b>Educational success</b>	<b>Cognitive development</b>	<b>Social-emotional development</b>	<b>Deviancy</b>
1. Special education 2. Graduation 3. Drop-out 4. Grade retention 5. Completed years of education 6. Absenteeism 7. Reduced learning problems 8. Feeling of belonging at school	1. IQ 2. Achievement tests 3. School grades 4. Rating of academic skill and performance 5. School failure	1. Self-esteem 2. Self-confidence 3. Social skills 4. Parent- teacher rating of problem behaviour 5. Obsessive-compulsive behaviour	1. Delinquent behaviour 2. Running away from home 3. Drug use 4. Lying about age 5. Gang involvement
<b>Social participation</b>	<b>Criminal justice outcomes</b>	<b>Family wellbeing</b>	
1. Employment in teen years 2. Socioeconomic success 3. Social integration 4. Engaged in skilled labour	1. Rates of juvenile arrest 2. Incarceration 3. Rates of violent arrest 4. Rates of non-violent arrest 5. Petition requests to juvenile courts	1. Child maltreatment 2. Parental-adolescent relationship 3. Family functioning 4. Parental involvement in child's schooling 5. Parental conflict	

### 9.3.2: Step 2 - Framework for selection and data collection

#### (a) Selecting participants

Participants ( $n = 25$  (100 percent response rate)) have been selected on the basis of (a) their ability to influence decisions regarding the implementation of

program options available for early childhood intervention; and/or (b) their demonstrated expertise and experience in evaluation of the effectiveness of outcomes (both short- and long-term) associated with existing early childhood intervention programs adopted in the Queensland context.

The selected participants come from four distinct stakeholder groups: a policy development group (e.g. representatives of Queensland Department of Communities, Department of Child Safety, Queensland Health, Department of Education, Training and the Arts); a school level group (e.g. school teachers and principals, co-ordinators of childcare centres, and co-ordinators of crèche and kindergartens); a community agencies group (e.g. management and senior staff of private community organisations involved in the delivery of community-based developmental intervention programs - for example, Mission Australia); and, an academic group (e.g. academic researchers who contribute to the literature relating to developmental prevention and early education).

***(b) Data collection***

Data collection in this research is based on a multiple informant model. Collecting reports from multiple informants provides a more accurate and reliable data set, particularly when data relates to the outcomes for children. This is particularly salient when assessing psychopathology, when the trait of interest is child/adolescent behavioural and emotional problems (Achenbach, McConaughy, & Howell, 1987). Researchers in psychiatric epidemiology frequently use this method given that children are often too young to provide reliable information on their

cognitive state, social-emotional state and personal behaviour (Breton, Bergeron, Valla, Lepine, Houde, & Gaudet, 1995; Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985). The assessment of psychopathology is typically error prone. Horton, Laird and Zahner (1999) state: "...most types of psychopathology are not easily described or classified into diagnostic categories...thus there is a lack of reproducibility in classification...these assessment problems are particularly pronounced for research involving children (p. 7)."

Data were collected by conducting two different surveys. The first survey related to the strength of the various outcomes (educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes, and family wellbeing) with respect to their potential contribution to increased non-health-related quality of life during the adolescent life years. Preferences among the outcomes were determined by developing a matrix that compared the various outcomes in pairs with respect to potential contribution to NHRQOL during adolescence. Moreover, we asked participants to express preferences among the intensities of the attributes by developing seven matrices that compare outcome levels of success (no effect, small effect, medium effect, high effect, and very high effect) in pairs with respect to each attribute. The overall goal was to develop a weighted vector of priorities for program desirability with respect to the most desired attribute intensities. The second survey aimed to determine perceived program standings (structured preschool program, home visitation, centre-based childcare/developmental day care, family support services, and parental education) in pairs with respect to the most desired attribute intensities derived from the first survey. A detailed explanation of this method is provided in Section 9.4.

Members of the policy development group and the academic group were selected to complete the first survey. These groups were chosen as it was envisaged that they could make the most informed choices based on their expert knowledge of adolescent outcomes associated with early childhood intervention programs. Also, it was recognised that this group could make expert choices based on their ability to understand objective data results from longitudinal research. The experts were presented with data which demonstrated the effects of early childhood intervention programs on the seven outcome domains during the adolescent life phase identified in Table 9.1. Data were presented from results of the meta-analysis (Study 1). The policy development and academic groups were asked to complete survey questions that related to the above-mentioned adolescent outcome domains.

The second survey was administered to both the school-level group and the community agencies group. The school-level and community agency groups were asked to complete survey questions that related to the five forms of early childhood intervention programs discussed above.

The surveys were split to ensure that questions posed to participants were directed toward their strengths in terms of knowledge and experience. Additionally, the survey instrument, in its entirety, would have been too large and complicated for a single respondent to answer in one sitting. Further, such a method may have produced intransitive responses as the respondent became tired. Consequently, it was decided that two survey instruments would be appropriate in this situation. Would

overall priorities have differed had respondents been asked to answer all questions (assuming this was possible) is speculative, and would need to be tested.

*(c) Sources and methods of survey collection*

Prior to the implementation of the survey, an information letter together with a participant consent form was sent via email to all potential participants. The letter contained information such as: a description of the project; information on what the project aimed to achieve; a statement highlighting what information would be collected; a description of the expected benefits and risks to the participant; issues regarding confidentiality; a statement regarding voluntary participation; and methods of contacting researcher/s for further information. This letter is reproduced as Appendix E.

To confirm receipt of the letter and consent form, a phone call was made to all potential participants. In this call, potential participants were asked to indicate their willingness to be involved in the project. For those willing to be involved, a date, time, and location were negotiated. It was explained that the survey would take approximately 30-40 minutes to complete.

Methods used to collect data from the various sources comprised:

- Policy development interviews ( $n = 5$ ). Interviews were conducted with representatives of the above mentioned government departments (based on their individual perspective, not the perspective of their respective government departments). The survey was conducted face-to-face with the

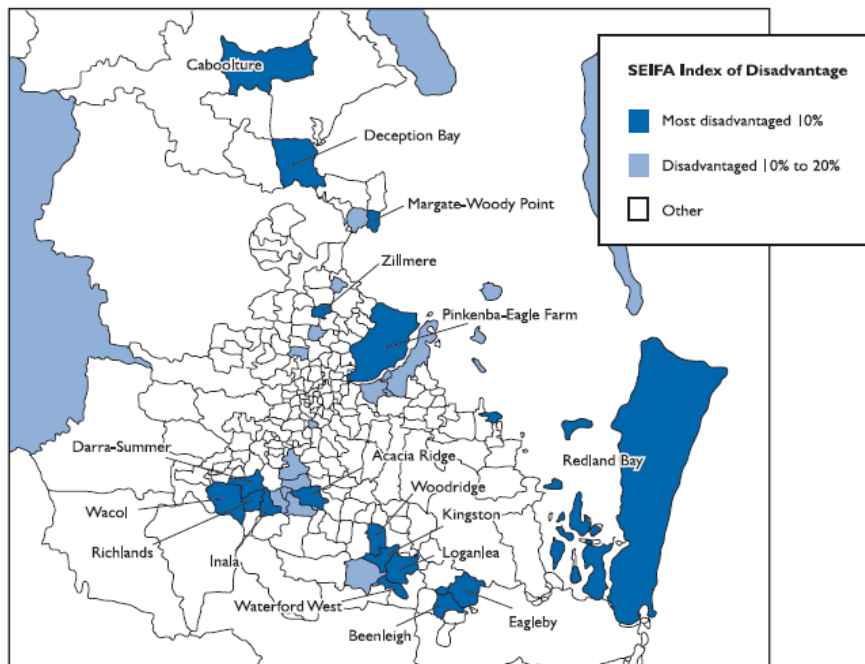


participants. This was seen as the most viable option given that objective data results had to be presented to participants. Prior to the interview, participants were sent a small discussion paper regarding the effects of early childhood intervention on non health-related outcomes during adolescence. This discussion paper is reproduced as Appendix F. Before beginning the interview, the participant was once again asked whether he/she had read the information letter and agreed to participate in the survey. Implied consent was given as a result of the participants completing the data collection instrument. Further, a coversheet which communicated the required information/assurances was attached to the survey and participants retained this coversheet. This data collection method followed the procedures and guidelines set out in Booklet 22: Informed consent procedures - section 5: Other informed consent procedures (Office for Research Griffith University, 2007), and was in accordance with the approved ethics application for this project (see Appendix G for a copy of this approval).

- School group interviews ( $n = 8$ ). Interviews were conducted with the principals of Queensland State primary schools and principals/directors of learning and behavioural management schools (e.g. Tennyson Special School). To be eligible for inclusion, the principals of schools must have been currently working in, or had previously worked in, a school located in the top 10 percent of disadvantaged areas in the Greater Brisbane area. These areas were chosen given the strong evidence suggesting that children raised in socio-economically disadvantaged regions have an increased probability of negative outcomes such as school failure, delinquency, drug

abuse, juvenile crime, youth unemployment and teenage pregnancies (Farrington & Welsh, 2003; Karoly et al., 1998; Keating & Hertzman, 1999). Areas were selected based on data from the 1996 Index of Relative Socio-Economic Disadvantage (SEIFA) (Queensland Council of Social Services, 1999). Some of the areas include: Acacia Ridge, Beenleigh, Darra, Eagleby, Inala, Sumner Park, and Wacol (see Figure 9.2). A snapshot of one of the areas, namely the Inala community is provided in Table 9.2. This table demonstrates that the Inala community has significantly higher rates of single parent families, a greater percentage of people living in public housing commission properties, higher rates of unemployment, a significantly greater number of families living below the poverty line, and a higher concentration of people speaking a language other than English when compared to the calculated rates across the greater Brisbane area (Homel et al., 2006). The method adopted to conduct the school group interviews was identical to that used in the policy development interviews.

**Figure 9.2: Disadvantaged areas in Brisbane (1996 Index of Relative Socio-Economic Disadvantage SEIFA)**



Source: Homel, R., Freiberg, K., Lamb, C., Leech, M., Carr, A., Hampshire, A., Hay, I., Elias, G., Manning, M., Teague, R., & Batchelor, S. (2006). *The Pathways to Prevention Project: The first five years 1999-2004*. Brisbane: Griffith University and Mission Australia.

**Table 9.2: Snapshot of the Inala community taken from the 2001 Census**

2001 Census Data	Inala	Brisbane
Children aged 0-9 years	20.8%	13.8%
0-5 years	10.1%	8.1%
% single parent families	33.1	15.7
% public housing commission properties	39.1	4.3
% unemployed	20.0	7.8
% families below poverty line	40.7	18.9
Median weekly household income	\$400-499	\$800-899
% adult population who completed high school	23.6	43.2
% language other than English	32.3	10.1
% Indigenous	7.2	1.7
Rate of Court appearances in 1998 and 1999 (per 1000 10-16 year olds)	158	20

Source: Homel, R., Freiberg, K., Lamb, C., Leech, M., Carr, A., Hampshire, A., Hay, I., Elias, G., Manning, M., Teague, R., & Batchelor, S. (2006). *The Pathways to Prevention Project: The first five years 1999-2004*. Brisbane: Griffith University and Mission Australia.

- Community agencies group interviews ( $n = 7$ ). Interviews were conducted with staff associated with the implementation of the *Pathways to Prevention* Project. *Pathways to Prevention* was established in 2001 as an early intervention project in the most disadvantaged urban area of Brisbane (Queensland Council of Social Services, 1999). The Pathways project aims to create opportunities for positive change in developmental pathways of individuals that could potentially, without intervention, lead to an array of damaging outcomes for young people (e.g. school failure, drug abuse and juvenile crime). By working in partnership with their families and schools within a community development framework, *Pathways* attempts to influence social capital and improve outcomes for children (Homel et al., 2006; Manning, 2004). *Pathways* provides culturally appropriate services to children and their families who are disadvantaged due to barriers (e.g. healthy development and family functioning) caused by limited access to economic, personal and social resources (Freiberg, 2005; Homel et al., 2006). With the aim of enhancing individual and family wellbeing, *Pathways* integrates family-support and school-based programs to assist families to become engaged in their child's education and community life. This is achieved through the development of activities which develop networks that support a mechanism for creating strong links between individuals and institutions (Homel et al., 2006). The method adopted to conduct the community agencies group interviews was identical to that used in the policy development and school-level interviews.
- Academic group interviews ( $n = 5$ ). This group includes academic researchers and PhD students who contribute to the literature relating to developmental

prevention and early education. Participants were chosen from the School of Criminology and Criminal Justice, the Key Centre for Ethics, Law, Justice and Governance, the School of Education at Griffith University, and the School of Social Science at University of Queensland. Procedures used to conduct the academic group interviews data were identical to those employed with the other groups.

***(d) Participant confidentiality***

Four safeguards were incorporated to protect the privacy and confidentiality of the survey participants:

1. Signed and informed consent to collect and analyse the data was obtained;
2. All participants were advised that they could withdraw their consent for data to be analysed at any time;
3. All surveys are held in a locked filing cabinet, located within a locked office on university premises; and
4. All data has been aggregated in such a way that individual views or preferences cannot be identified.

***(e) Communication of results***

The responses from all surveys collected were aggregated, whereby a mathematical method was used to synthesise judgements and calculate relative weights of alternatives to obtain overall priorities indicating program desirability. Aggregated results were provided to interested participants. All participants were informed of their rights to view aggregated results. All aggregated result data is

anonymous and cannot be linked to individual participants. Also, the names of those who participated in the survey will not be published or released to those outside the research team. Only aggregated data were incorporated into the results of this thesis.

### **9.3.3: Step 3 - Scenario of associated outcomes**

Before elements of each level of the hierarchy were rated using pair-wise comparisons, a small discussion paper highlighting the effects of early childhood intervention on at-risk families and their children on non health-related outcomes during adolescence was presented to participants. Information included: a definition on non health-related quality of life; factors that contribute to non health-related quality of life; a description of the five forms of early childhood intervention used in the analysis a definition of outcome domains and their indicators; a definition and description of effect sizes and what they mean; and, a diagram demonstrating the effects of early childhood intervention programs on outcomes during adolescence (see Appendix F). The aim of this paper was to provide participants with information regarding the effectiveness of early childhood intervention on non health-related outcomes during adolescence, so that they could make more informed choices when making pair-wise comparisons between the various elements within the hierarchy. The outcome data incorporated in this information paper was derived from the meta-analysis reported in Chapter 6.

### **9.3.4: Step 4 - Setting up the decision hierarchy**

The decision hierarchy developed for this study is given in Figure 9.3 and states a desired future outcome or goal (program desirability with respect to its

perceived contribution to non-health related quality of life during the adolescent life years), then through the use of a backward process determines the actions that are needed to achieve the desired outcome (Alexander & Saaty, 1977b). The backward process, considered a normative process, can be seen as a policy design vehicle for determining the means to bring about desired future outcome/s. This is a highly effective vehicle given that it allows policy makers at all levels to analyse complex, multi-criteria problems. Through its focus on trade-offs, it helps to identify the advantages and disadvantages of policy options under the veil of uncertainty (Saaty, 1994).

Two assumptions are made in the formation of the hierarchy, without which the problem cannot be analysed using the AHP method: (1) each element of a level in a hierarchy is related to other elements in adjacent levels, and (2) no relationship exists between elements on the same level (Cheng & Li, 2001a; Saaty, 1990).

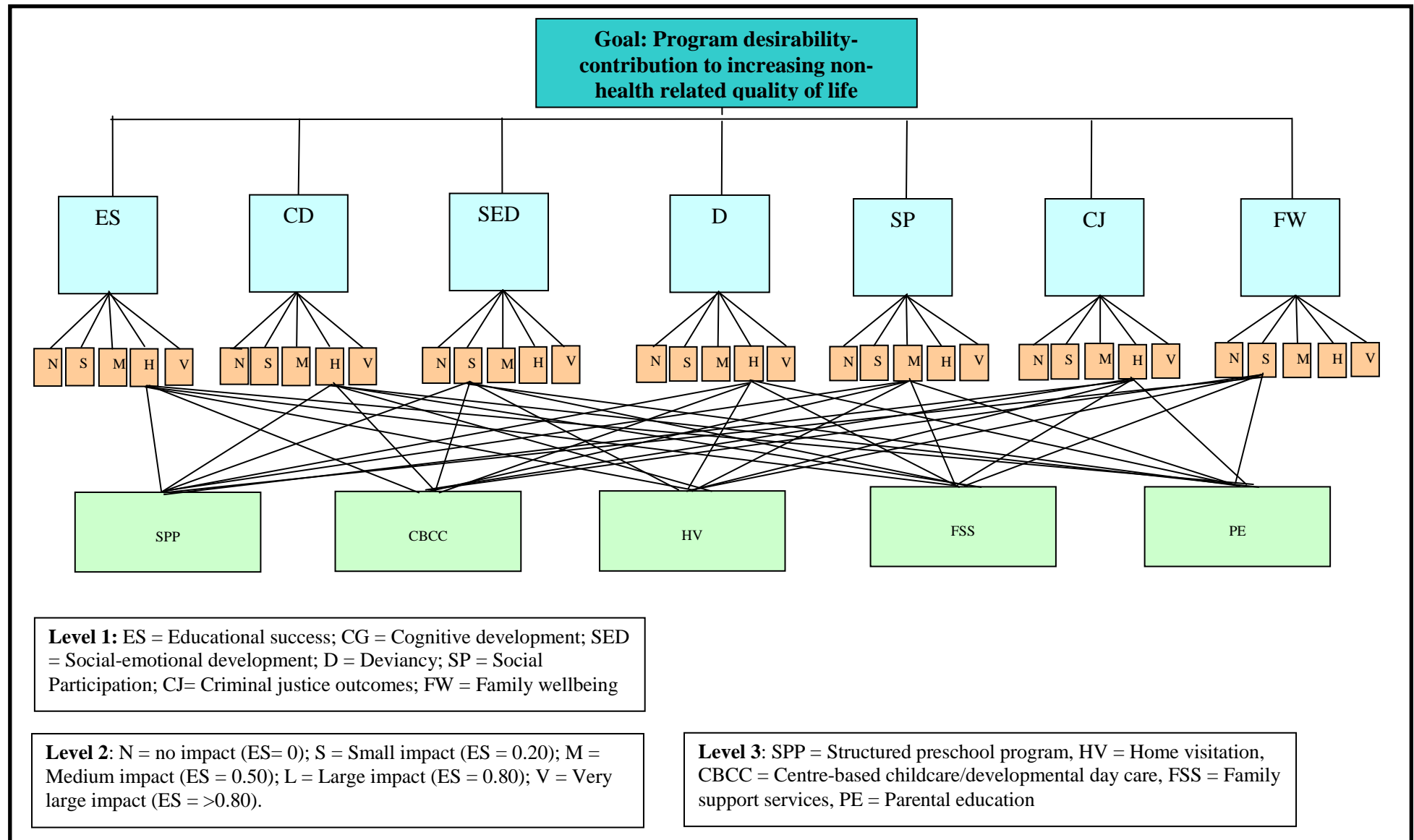
The top of the hierarchy represents the overall goal (program desirability based on contribution to increasing non-health related quality of life during the adolescent life phase). The next level down (Level 1) represents attributes considered most relevant to achieving improvements in non-health related quality of life. Rather than relying on subjective perceptions of what attributes should be considered (ratings from various individuals), a better method has been adopted in this study whereby attributes have been derived from objective research (specifically, the meta-analysis reported in Chapter 6). This includes weightings of non health-related outcome success of early childhood interventions on each of the seven attributes or domains (educational success; cognitive development; social-emotional development;

deviancy; social participation; criminal justice outcomes; and, family wellbeing). Level 2 of the hierarchy highlights five possible outcomes (no impact, small impact, medium impact, large impact and very large impact), which may result from the various attributes or domains in level 1 of the hierarchy. These intensity levels are derived from Cohen (1992). Level 3 provides the various preschool program options that respondents are asked to consider. Essentially, this level of the hierarchy provides five options of preschool intervention programs (structured preschool program, centre-based childcare/developmental day care, home visitation, family support services, and parental education) that potentially contribute to an increase in non-health-related quality of life of an individual during the adolescent life phase.

The benefit of this structure is that it allows us to make judgements on the relative importance of each outcome domain, the relative importance of their determinants/indicators and ultimately to determine the relative importance or desirability of each possible program option.



**Figure 9.3: Decision hierarchy for early childhood intervention program alternatives, outcome domains and level of effect**



### 9.3.5: Step 5 - Pair-wise comparisons

Saaty's (1990) scale of measurement used to rate the intensity of importance between two elements is employed in this study (see Table 8.1). A semantic scale of 1 to 9 is adopted whereby individuals rate their perceived importance of one criterion over another. A detailed discussion of the use of this scale, together with an example, was provided in Section 8.4.

### 9.3.6: Step 6: Estimating relative weights

Preferences among the attributes were determined by developing a matrix that compared the various attributes in pairs with respect to contribution to non-health related quality of life during the adolescent life phase (CNHRQOL) (see Table 9.4). Results are based on responses from the survey. In this case, responses were provided by the policy and academic groups.

**Table 9.4: Matrix comparing outcome domains/attributes**

<b>CNHRQOL</b>	<b>ES</b> (Educational success)	<b>CD</b> (Cognitive development)	<b>SED</b> (Social-emotional development)	<b>D</b> (Deviancy)	<b>SP</b> (Social participation)	<b>CJ</b> (Criminal justice)	<b>FW</b> (Family wellbeing)
<b>ES</b>	$\alpha_{ESES}$	$\alpha_{ESCD}$	$\alpha_{ESSED}$	$\alpha_{ESD}$	$\alpha_{ESSP}$	$\alpha_{ESCJ}$	$\alpha_{ESFW}$
<b>CD</b>	$\alpha_{CDES}$	$\alpha_{CDCD}$	$\alpha_{CDSSED}$	$\alpha_{CDD}$	$\alpha_{CDSPP}$	$\alpha_{CDCJ}$	$\alpha_{CDFW}$
<b>SED</b>	$\alpha_{SEDES}$	$\alpha_{SEDCD}$	$\alpha_{SEDSSED}$	$\alpha_{SEDD}$	$\alpha_{SEDSP}$	$\alpha_{SEDCJ}$	$\alpha_{SEDFW}$
<b>D</b>	$\alpha_{DES}$	$\alpha_{DCD}$	$\alpha_{DSED}$	$\alpha_{DD}$	$\alpha_{DSP}$	$\alpha_{DCJ}$	$\alpha_{DFW}$
<b>SP</b>	$\alpha_{SPES}$	$\alpha_{SPCD}$	$\alpha_{SPSED}$	$\alpha_{SPD}$	$\alpha_{SPSP}$	$\alpha_{SPCJ}$	$\alpha_{SPFW}$
<b>CJ</b>	$\alpha_{CJES}$	$\alpha_{CJCD}$	$\alpha_{CJSED}$	$\alpha_{CJD}$	$\alpha_{CJSP}$	$\alpha_{CJCJ}$	$\alpha_{CJFW}$
<b>FW</b>	$\alpha_{FWES}$	$\alpha_{FWCD}$	$\alpha_{FWSED}$	$\alpha_{FWD}$	$\alpha_{FWSP}$	$\alpha_{FWCJ}$	$\alpha_{FWFW}$
<b>Total</b>	<b>*<math>\alpha_{ES}</math></b>	<b>*<math>\alpha_{CD}</math></b>	<b>*<math>\alpha_{SED}</math></b>	<b>*<math>\alpha_D</math></b>	<b>*<math>\alpha_{SP}</math></b>	<b>*<math>\alpha_{CJ}</math></b>	<b>*<math>\alpha_{FW}</math></b>

Adapted from: Saaty, T. L. (2001). *Decision making for leaders* (3rd Ed.). Pittsburgh, PA: RWS Publications.

\*Note: an unweighted average scheme is applied with respect to aggregation of individual responses

After developing the pair-wise comparison matrix, a vector of priorities or eigenvector is calculated and normalised to sum to 1 or 100 percent. We obtain the eigenvector by adding the elements in each row and dividing the sum by the number of elements in the row to obtain the “priority” or “relative weight” of each element (Saaty, 1990, 1994). Section 8.2 provided a detailed example of calculation of a vector of priorities. The priority vectors of each of the matrices were derived by utilizing the Expert Choice software (Expert Choice, 2000-2004). However, to ensure that the software was accurate, all eigenvectors and priority rankings were also calculated using an Excel spreadsheet. Moreover, the Expert Choice software did not provide detailed information on eigenvalues or the consistency index; therefore it was necessary to use this supplementary method. Copies of the various Excel spreadsheets are available on request.

Next, preferences among the intensities of the attributes were determined by developing seven matrices (e.g.  $\beta$  for ES,  $\gamma$  for CD,  $\delta$  for SED,  $\epsilon$  for D,  $\zeta$  for SP,  $\eta$  for CJ, and  $\theta$  for FW) that compared outcome levels of success (no impact, small impact, medium impact, large impact, and very large impact) in pairs with respect to each attribute. Table 9.5 provides an example of one of the matrices.

**Table 9.5: Matrix comparing intensity levels with respect to the outcome/attribute educational success (ES)**

<b>ES</b>	<b>N</b> (no impact)	<b>S</b> (small impact)	<b>M</b> (medium impact)	<b>H</b> (high impact)	<b>V</b> (very high impact)	<b>Priority</b>
<b>N</b> (no impact)	$\beta_{NN}$	$\beta_{NS}$	$\beta_{NM}$	$\beta_{NH}$	$\beta_{NV}$	$*\beta_N$
<b>S</b> (small impact)	$\beta_{SN}$	$\beta_{SS}$	$\beta_{SM}$	$\beta_{SH}$	$\beta_{SV}$	$*\beta_S$
<b>M</b> (medium impact)	$\beta_{MN}$	$\beta_{MS}$	$\beta_{MM}$	$\beta_{MH}$	$\beta_{MV}$	$*\beta_M$
<b>H</b> (high impact)	$\beta_{HN}$	$\beta_{HS}$	$\beta_{HM}$	$\beta_{HH}$	$\beta_{HV}$	$*\beta_H$
<b>V</b> (very high impact)	$\beta_{VN}$	$\beta_{VS}$	$\beta_{VM}$	$\beta_{VH}$	$\beta_{VV}$	$*\beta_V$
<b>Total</b>						$*\beta$

Adapted from: Saaty, T. L. (2001). *Decision making for leaders* (3rd ed.). Pittsburg, PA: RWS Publications.

### 9.3.7: Step 7 - Calculating consistency to validate results

Given that most people are inconsistent in answering questions, it is necessary to measure the consistency of each derived vector by calculating a consistency ratio (CR) (Saaty, 2000). Essentially, a consistency ratio is used to measure the consistency of the pair-wise comparisons made by individual respondents (Isard & Smith, 1982). Saaty (1994) proposes a set of CR values that are acceptable for various matrix sizes. A CR value of 0.05 is argued to be acceptable for 3 x 3 matrixes, 0.08 is acceptable for a 4 x 4 matrix, and 0.1 is acceptable for larger matrices. A generated pair-wise comparison response CR that falls into an acceptable range can be included into the AHP process. A larger value would imply excessive intransitivity of preferences (Cheng & Li, 2001a). The procedure for calculating the CR was discussed in detail in Section 8.4.

When CR values fall within acceptable ranges, participant results are aggregated to obtain the combined judgements on the weight of elements at each level of the hierarchy. However, if CR values do not fall within acceptable ranges, then two methods for improving consistency can be considered. Method 1 involves asking participants with inconsistent preferences to provide another set of answers to the pair-wise comparisons. The second method requires a more radical approach whereby the problem is restructured evolving into a new grouping of elements under a more meaningful attribute schema. This requires the development of the hierarchy to construct alternative questions so that a more consistent set of answers may be sought. Finally, questionnaires that are incomplete should be removed from the sample. In this research, all complete surveys were retained, apart from those which were incomplete. These surveys (n=2) were subsequently removed from the sample.

#### **9.3.8: Step 8 - Synthesise judgements/ Calculate relative weights of alternatives to obtain overall priorities indicating program desirability**

Step seven involves synthesising judgements with the aim of obtaining a set of overall vector of priorities of preferences for the alternative programs. This vector indicates which early childhood intervention program participants consider more desirable in regards to contributing to non-health related quality of life during the adolescent life phase. To achieve this, the priorities of the five levels of intensity (N, S, M, H, V) are grouped for each of the six attributes, taken from the seven vectors (e.g.  $\beta$  for ES,  $\gamma$  for CD,  $\delta$  for SED,  $\epsilon$  for D,  $\zeta$  for SP,  $\eta$  for CJ, and  $\theta$  for FW) derived in step 6. This generates a matrix of the form given in Table 9.6.

**Table 9.6: Priorities of attributes/outcomes with respect to five levels of intensity**

		Attributes						
Levels of Intensities		ES	CD	SED	D	SP	CJ	FW
	N	$*\beta_N$	$*\gamma_N$	$*\delta_N$	$*\epsilon_N$	$*\zeta_N$	$*\eta_N$	$*\theta_N$
	S	$*\beta_S$	$*\gamma_S$	$*\delta_S$	$*\epsilon_S$	$*\zeta_S$	$*\eta_S$	$*\theta_S$
	M	$*\beta_M$	$*\gamma_M$	$*\delta_M$	$*\epsilon_M$	$*\zeta_M$	$*\eta_M$	$*\theta_M$
	H	$*\beta_H$	$*\gamma_H$	$*\delta_H$	$*\epsilon_H$	$*\zeta_H$	$*\eta_H$	$*\theta_H$
	V	$*\beta_V$	$*\gamma_V$	$*\delta_V$	$*\epsilon_V$	$*\zeta_V$	$*\eta_V$	$*\theta_V$

Next, each column in this matrix is multiplied by the priority of the corresponding attribute (obtained in step 6) to obtain weighted vectors of priority for all levels of intensity. For example, to obtain a weighted vector priority for cell (column 1, row 1) we have  $(*\beta_N) \times (*\alpha_{ES}) = \Omega_{NES}$ , for cell (column 1, row 2) we have  $(*\beta_S) \times (*\alpha_{ES}) = \Omega_{SES}$  and so on. The resulting matrix is given in Table 9.7.

**Table 9.7: Weighted vectors of priority for all levels of intensity**

		Attributes						
Attribute Intensities		ES	CD	SED	D	SP	CJ	FW
	N	$\Omega_{NES}$	$\Omega_{NCD}$	$\Omega_{NSED}$	$\Omega_{ND}$	$\Omega_{NSP}$	$\Omega_{NCJ}$	$\Omega_{NFW}$
	S	$\Omega_{SES}$	$\Omega_{SCD}$	$\Omega_{SSED}$	$\Omega_{SD}$	$\Omega_{SSP}$	$\Omega_{SCJ}$	$\Omega_{SFW}$
	M	$\Omega_{MES}$	$\Omega_{MCD}$	$\Omega_{MSED}$	$\Omega_{MD}$	$\Omega_{MSP}$	$\Omega_{MCJ}$	$\Omega_{MFW}$
	H	$\Omega_{HES}$	$\Omega_{HCD}$	$\Omega_{HSED}$	$\Omega_{HD}$	$\Omega_{HSP}$	$\Omega_{HCJ}$	$\Omega_{HFW}$
	V	$\Omega_{VES}$	$\Omega_{VCD}$	$\Omega_{VSED}$	$\Omega_{VD}$	$\Omega_{VSP}$	$\Omega_{VCJ}$	$\Omega_{VFW}$

We then select, from each column in Table 9.7, the element with the highest priority to obtain a vector of desired attribute intensities (e.g.  $\Omega_{HES}$ ,  $\Omega_{HCD}$ ,  $\Omega_{SSED}$ ,  $\Omega_{VD}$ ,  $\Omega_{MSP}$ ,  $\Omega_{HCJ}$ ,  $\Omega_{SFW}$ ). We then sum this vector ( $\Omega_{HES} + \Omega_{HCD} + \Omega_{SSED} + \Omega_{VD} + \Omega_{MSP} + \Omega_{HCJ} + \Omega_{SFW}$ ) to arrive at a total ( $*\Omega$ ). Next, we divide each entry (e.g.  $\Omega_{HES}$ ,  $\Omega_{HCD}$ ,  $\Omega_{SSED}$ ,  $\Omega_{VD}$ ,  $\Omega_{MSP}$ ,  $\Omega_{HCJ}$ ,  $\Omega_{SFW}$ ) by ( $*\Omega$ ) to get the normalised vector of desired attribute intensities (e.g.  $\Omega_{HES}/*\Omega = *\Omega_{HES}$ ,  $\Omega_{HCD}/*\Omega = *\Omega_{HCD}$ ). This vector is of the form given in Table 9.8.

**Table 9.8: Normalised vector of desired attribute/outcome intensities**

Vector of desired attribute intensities
$*\Omega_{HES}$
$*\Omega_{HCD}$
$*\Omega_{SSED}$
$*\Omega_{VD}$
$*\Omega_{MSP}$
$*\Omega_{HCJ}$
$*\Omega_{SFW}$

We then develop matrices for each intensity-attribute comparison that compare the five preschool program alternatives (SPP = Structured preschool program, HV = Home visitation, CBCC = Centre-based childcare/developmental day care, FSS = Family support services, PE = Parental education) with respect to the most desired attribute intensities. Table 9.9 provides an example of one of these matrices developed for the H-ES ( $*\Omega_{HES}$ ) combination.

**Table 9.9: Matrix comparing the five early childhood intervention program alternatives for desired attribute/outcome intensities**

$*\Omega_{HES}$	Attribute Intensities					
	SPP	CBCC	HV	FSS	PE	Priority
SPP	$\beta_{SPP-SPP}$	$\beta_{SPP-CBCC}$	$\beta_{SPP-HV}$	$\beta_{SPP-FSS}$	$\beta_{SPP-PE}$	$*\beta_{SPP}$
CBCC	$\beta_{CBCC-SPP}$	$\beta_{CBCC-CBCC}$	$\beta_{CBCC-HV}$	$\beta_{CBCC-FSS}$	$\beta_{CBCC-PE}$	$*\beta_{CBCC}$
HV	$\beta_{HV-SPP}$	$\beta_{HV-CBCC}$	$\beta_{HV-HV}$	$\beta_{HV-FSS}$	$\beta_{HV-PE}$	$*\beta_{HV}$
FSS	$\beta_{FSS-SPP}$	$\beta_{FSS-CBCC}$	$\beta_{FSS-HV}$	$\beta_{FSS-FSS}$	$\beta_{FSS-PE}$	$*\beta_{FSS}$
PE	$\beta_{PE-SPP}$	$\beta_{PE-CBCC}$	$\beta_{PE-HV}$	$\beta_{PE-FSS}$	$\beta_{PE-PE}$	$*\beta_{PE}$

We then group the priorities of the five early childhood intervention program alternatives with respect to each desired attribute-intensity in columns, and enter the priorities for the vector of desired attribute-intensities (from Table 9.8) above the

columns. We then multiply each column (in Table 9.9) by the normalised priority of the corresponding attribute-intensity to obtain the weighted vectors of priority for the desired attribute-intensities for each early childhood intervention program alternative. For example, to obtain a weighted vector priority for cell (column 1, row 1) we have  $(^*\beta_{SPP}) \times (^*\Omega_{HES}) = \Psi_{SPP/H-ES}$ , for cell (column 1, row 2) we have  $(^*\beta_{CBCC}) \times (^*\Omega_{HES}) = \Psi_{CBCC/H-ES}$ . The resulting matrix is given in Table 9.10.

**Table 9.10: Weighted vector of priorities for desired attribute/outcome intensities for early childhood intervention alternatives**

	$(^*\Omega_{HES})$ <b>H-ES</b>	$(^*\Omega_{HCD})$ <b>H-CD</b>	$(^*\Omega_{SSED})$ <b>S-SED</b>	$(^*\Omega_{VD})$ <b>V-D</b>	$(^*\Omega_{MSP})$ <b>M-SP</b>	$(^*\Omega_{HCJ})$ <b>H-CJ</b>	$(^*\Omega_{SFW})$ <b>S-FO</b>
<b>SPP</b>	$\Psi_{SPP/H-ES}$	$\Psi_{SPP/H-CD}$	$\Psi_{SPP/S-SED}$	$\Psi_{SPP/V-D}$	$\Psi_{SPP/M-SP}$	$\Psi_{SPP/H-CJ}$	$\Psi_{SPP/S-FW}$
<b>CBCC</b>	$\Psi_{CBCC/H-ES}$	$\Psi_{CBCC/H-CD}$	$\Psi_{CBCC/S-SED}$	$\Psi_{CBCC/V-D}$	$\Psi_{CBCC/M-SP}$	$\Psi_{CBCC/H-CJ}$	$\Psi_{CBCC/S-FW}$
<b>HV</b>	$\Psi_{HV/H-ES}$	$\Psi_{HV/H-CD}$	$\Psi_{HV/S-SED}$	$\Psi_{HV/V-D}$	$\Psi_{HV/M-SP}$	$\Psi_{HV/H-CJ}$	$\Psi_{HV/S-FW}$
<b>FSS</b>	$\Psi_{FSS/H-ES}$	$\Psi_{FSS/H-CD}$	$\Psi_{FSS/S-SED}$	$\Psi_{FSS/V-D}$	$\Psi_{FSS/M-SP}$	$\Psi_{FSS/H-CJ}$	$\Psi_{FSS/S-FW}$
<b>PE</b>	$\Psi_{PE/H-ES}$	$\Psi_{PE/H-CD}$	$\Psi_{PE/S-SED}$	$\Psi_{PE/V-D}$	$\Psi_{PE/M-SP}$	$\Psi_{PE/H-CJ}$	$\Psi_{PE/S-FW}$

Finally, we added each of the five rows of Table 9.10 to obtain the overall vector of priorities for the five preschool program options. For example, the element in row 1 in this vector is calculated as:  $\Psi_{SPP/H-ES} + \Psi_{SPP/H-CD} + \dots + \Psi_{SPP/S-FW} = ^*\Psi_{SPP}$ . The complete set of row entries in vector  $\Psi$  are provided in Table 9.11.

**Table 9.11: Overall weighted vector of priorities for the five early childhood intervention program options**

<b>Vector of desired attribute intensities</b>
$^*\Psi_{SPP}$
$^*\Psi_{CBCC}$
$^*\Psi_{HV}$
$^*\Psi_{FSS}$
$^*\Psi_{PE}$



From the results obtained in Table 9.11 we can identify the program with the highest priority (i.e. the one with the highest numerical value) as the most desired preschool program in regards to its contribution to non-health related quality of life during the adolescent life phase.

#### **9.4: Conclusion**

In this chapter we examined the need to develop a tool for making complex, multi-criteria decisions that affect policy with respect to early childhood/early-in-life crime prevention programs. We posit that although policy decisions are often made with consideration of empirical research, complex decisions of this nature cannot be made without a structured methodological process. Through the use of the AHP method, policy-makers are able to solve complex problems and make better decisions regarding potential funding based on priorities assigned to attributes (e.g. educational success, cognitive development), objectives (e.g. levels of success), and the various alternatives (e.g. structured preschool programs, home visitation etc.). Further, the AHP method captures preference values for health and non health-related outcomes so that economists can measure the cost-utility of prevention programs. Moreover, the method provides a set of common metric outcomes so that seemingly disparate early-in-life prevention programs can be compared.

We have also described in detail the AHP method adopted in this thesis for determining preferences values with respect to non health-related outcomes associated with early childhood interventions for at-risk families and their children during the

adolescent life phase. Chapter 10 provides the results of the survey of this method. Based on participant responses, we provide results regarding preference values for non health-related outcomes during adolescence, and vectors of priority for all levels of our hierarchy. Furthermore, we highlight the most preferred early childhood intervention program type.

## **CHAPTER 10: RESULTS FROM APPLICATION OF THE ANALYTIC HIERARCHY PROCESS TO EVALUATION OF ALTERNATIVE EARLY INTERVENTION PROGRAMS**

### **10.1: Introduction**

Chapter 10 provides the results of our application of the analytic hierarchy process (AHP) to the evaluation of alternative early intervention programs. These results include preference values obtained by using the AHP method. Preference values, which as discussed in Chapter 7, may be used to measure the economic impact of qualitative improvements in non health-related quality of life (during adolescence) resulting from early childhood intervention programs. Section 10.2 provides results from our first survey. These results include preference values for seven outcomes during the adolescent life phase, preferences among the intensities of these outcomes, and pair-wise comparisons of five levels of success (no effect, small effect, medium effect, high effect, and very high effect) with respect to each outcome (ES, CD, SED, D, SP, CJ and FW). Section 10.3 presents the results of our second survey. These results include pair-wise comparisons of the five preschool program options (structured preschool program (SPP), centre-based developmental day care (CBCC), home visitation (HV), family support services (FSS), and parental education (PE)), with respect to the most desired attribute outcome/level of success (e.g. VH-ES, VH-CD, and VH-SED). Section 10.4 aggregates and summarises results from surveys 1 and 2 to generate overall priorities or preference rankings for the five early childhood intervention program alternatives with respect to their perceived impact on non-health related quality of life during adolescence. In section 10.5, the stability of the results is tested using a sensitivity analysis. The results are

discussed in more detail in Section 10.6, and the limitations of the study are identified in Section 10.7. Finally, in Section 10.8, a brief conclusion is provided and the final chapter of the thesis is introduced.

## **10.2: Results of survey 1**

Survey 1 was comprised of two sections. The first section of the survey was designed to provide insight into the relative preference and strength of seven outcomes (educational success (ES), cognitive development (CD), social-emotional development (SED), deviancy (D), social participation (SP), criminal justice outcomes (CJ), and family wellbeing (FW)) with respect to their potential contribution to increasing non-health related quality of life (NHRQOL) during adolescence. In the second section, a further seven matrices were developed to determine preferences among the intensities of the outcomes, involving pair-wise comparisons of five levels of success (no effect, small effect, medium effect, high effect, and very high effect) in pairs with respect to each outcome. Members of the policy development group (n=5) and the academic group (n=5) were selected to complete this first survey. Sections 1 and 2 of the first survey comprised a total of 91 questions, which took approximately 30 - 40 minutes for respondents to complete. A copy of this first survey is provided in Appendix H.

### **10.2.1: Results of survey 1 section 1**

Section 1 of the first survey comprised 21 questions relating to the relative preference and strength of seven outcomes (ES, CD, SED, D, SP, CJ, and FW) with respect to their potential contribution to increasing non-health related quality of life

(NHRQOL) during adolescence. Results from the 10 respondents were aggregated and the average of these responses identified. From these results Table 10.1 and Figure 10.1 highlight the relative preference, strength and priorities of seven outcomes domains with respect to their potential contribution to increasing non-health related quality of life (NHRQOL) during adolescence. Appendix I provides descriptive statistics for all 21 questions comprising this section of the survey.

**Table 10.1: Pair-wise comparison of adolescent outcomes (average response)**

	ES	CD	SED	D	SP	CJ	FW	
ES	1.00	0.79	0.30	0.52	1.08	0.89	0.28	<div> <div>Priority</div> <div>0.078</div> <div>0.063</div> <div>0.161</div> <div>0.113</div> <div>0.144</div> <div>0.112</div> <div>0.330</div> </div>
CD	1.27 (1.42)	1.00	0.28	0.42	0.58	0.39	0.26	
SED	3.33 (2.36)	3.55 (2.43)	1.00	0.84	0.70	1.02	0.78	
D	1.92 (2.44)	2.36 (3.10)	1.19 (1.90)	1.00	0.37	0.99	0.23	
SP	0.93 (1.46)	1.73 (1.90)	1.43 (1.73)	2.71 (2.38)	1.00	1.19	0.28	
CJ	1.12 (1.92)	2.54 (2.50)	0.98 (1.20)	1.01 (1.11)	0.84 (1.14)	1.00	0.29	
FW	3.63 (2.63)	3.90 (1.60)	1.27 (1.05)	4.43 (2.09)	3.55 (2.34)	3.46 (2.42)	1.00	

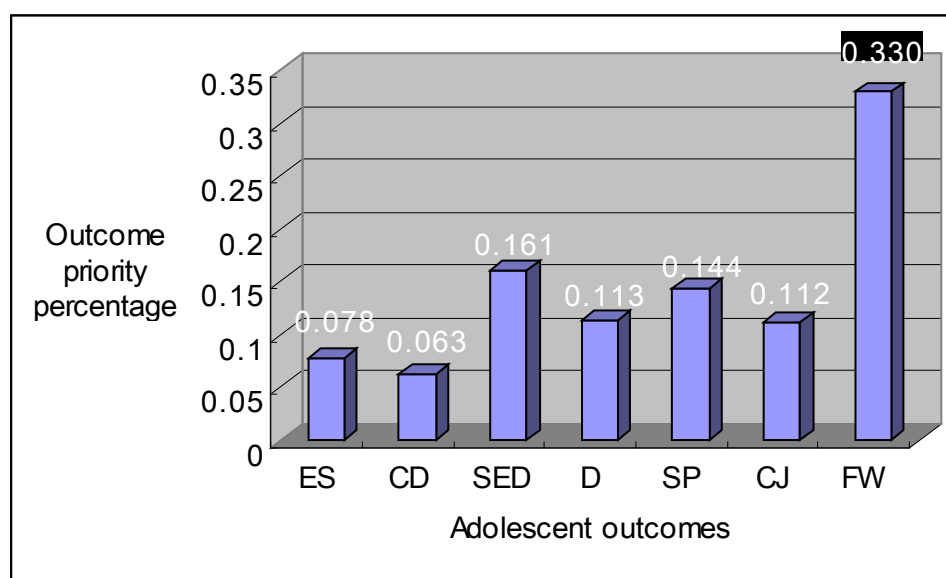
$\lambda_{\max}=7.48$ ; C.I = 0.08; C.R = 0.06

Note: figures in brackets represent the standard deviation across 10 respondents; Shaded figures represent the reciprocal of average responses

Results from Table 10.1 and Figure 10.1 demonstrate that the outcome family wellbeing (FW) is considered the highest priority during adolescence (0.330) followed by social-emotional development (SED) (0.161). Entries in Table 10.1 range from 1.00 (equal importance) to 4.33 (slightly less than strong importance of one element over another (see Table 8.1)). Although Saaty's scale ranges from 1 to 9, it seems plausible to find an average of responses with an upper limit of 4 to 5. Suggesting that respondents do not consider any of the alternatives to be of very strong, or of absolute importance over

another. Additionally, Appendix I demonstrates that there is a tendency for the standard deviation of some pairwise comparisons to increase as the mean increases. This could possibly reflect increased disagreement among participants with respect to some pairwise comparisons between the various elements. However, this is not common, and overall the consistency of responses ( $C.I. = 0.08$ ) reflect consistency among survey respondents.

**Figure 10.1: Priority percentage for outcomes during adolescence**



Social participation (SP) (0.144), criminal justice outcomes (CJ) (0.112), deviancy (D) (0.113), educational success (ES) (0.078) and cognitive development (CD) (0.06) constitute the remaining outcomes in order of preference (Figure 10.1). This is a particularly interesting result given that a meta-analytic review of outcomes associated with early childhood intervention (see Chapter 6) demonstrated that the mean weighted effect size ( $d$ ) for outcome domains CJ and FW are 0.244 and 0.204 respectively, highlighting small effects. Mean effect sizes for the outcome domains CD, SP and D were 0.334, 0.373 and 0.39 respectively, which indicates a small to medium effect.

Moreover, the largest effect size is displayed on the domain ES (0.532), which equates to a medium to medium-high effect, while the SED domain reveals the smallest effect (0.148). Results of the meta-analytic review were provided to all respondents prior to them responding to the survey; however, our analysis reveals that objective research results do not appear to determine perceived preferences for the relative importance of outcomes to NHRQOL during adolescence. Moreover, what respondents consider to be a high priority during adolescence (e.g. family wellbeing and social-emotional development) with respect to contributing to NHRQOL, is demonstrated by objective research to have relatively small effect sizes in terms of the effect of early childhood interventions on outcomes during adolescence (e.g. family wellbeing  $d = 0.204$  and social-emotional development  $d = 0.148$ ).

### **10.2.2: Results of survey 1 section 2**

Section 2 of the first survey comprised 70 questions relating to preferences among the intensities of the outcomes. Seven matrices were developed to provide insights into preferences among the intensities of the outcomes, comparing five outcome levels of success (no effect, small effect, medium effect, high effect, and very high effect) in pairs with respect to each of 7 outcomes (ES, CD, SED, D, SP, CJ and FW). Once again results from the 10 respondents were aggregated and the average of these responses identified for each of the above-mentioned matrices in turn. Tables 10.2–10.8 highlight the relative preference, strength and priorities among the intensities of the outcomes, comparing five outcome levels of success (no effect, small effect, medium effect, high effect, and very

high effect) in pairs with respect to each outcome. Appendix I provides descriptive statistics for all 70 questions comprising this section of the survey.

Tables 10.2 -10.8 highlight that a very high (VH) outcome level of success is preferred to all other levels of success. On average, this outcome level of success contributed to approximately 40 percent of total priority across all outcomes. However, on the outcome domain FW, the level of success VH contributed to only 36 percent of the total priority across the five levels of success. Very high (VH) and high (H) levels of success contributed to approximately 69 percent of the total priority across all outcomes. Further, outcome levels of success no effect (N) (0.04), small (S) (0.09), and medium (M) (0.18) make up approximately 31 percent of the remaining priority across all outcomes.

Table 10.2 highlights the outcome level of success with respect to the domain educational success (ES), and reveals that the levels of success very high (VH) and high (H) were considered a strong importance compared to no effect (N). This is indicated by an average priority rating of (6.3) (definitions of importance ratings are provided in Table 9.3). A medium effect (M) was considered a strong importance (4.8) and small effects (S) were considered of weak importance (3.2) when compared to no effect.



**Table 10.2: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome educational success (ES)**

ES	N	S	M	H	VH
N	1.00	0.31	0.21	0.16	0.16
S	3.20	1.00	0.30	0.22	0.23
M	4.80	3.30	1.00	0.41	0.34
H	6.20	4.60	2.43	1.00	0.55
VH	6.40	4.38	2.92	1.82	1.00

→	Priority
	0.04
	0.09
	0.18
	0.29
	0.40

$$\lambda_{\max} = 5.19; C.I = 0.05; C.R = 0.04$$

Further, VH and H levels of success when compared to a small effect were approximately 4.49, indicating an intermediate ranking between weak and strong importance. Moreover, a high effect was given a slightly higher preference score (4.6) compared to a very high effect (4.38). Additionally, a medium effect (M) was considered of weak importance over a small effect (S) (3.30). A VH effect was considered slightly less than of weak importance when compared to an M effect (2.92); while, a high effect (H) was also considered slightly less than of weak importance when compared to an M effect (2.43). Finally, a VH effect was considered only slightly more important than an H effect (1.82).

Table 10.3 compares outcome levels of success with respect to the domain cognitive development, indicate that the levels of success VH and H are considered of strong importance when compared to no effect (average 5.03); while a medium effect was slightly lower highlighting an intermediate ranking between weak and strong importance (4.25) when compared to no effect. A small effect (S) also highlighted an intermediate ranking between weak and strong importance (3.40) when compared to no effect.

**Table 10.3: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome cognitive development (CD)**

CD	N	S	M	H	VH
N	1.00	0.29	0.24	0.20	0.20
S	3.40	1.00	0.34	0.26	0.25
M	4.25	2.90	1.00	0.38	0.38
H	4.93	3.85	2.60	1.00	0.53
VH	5.13	4.07	2.62	1.90	1.00

$\rightarrow$ 

$\left[ \begin{array}{c} \text{Priority} \\ 0.05 \\ 0.09 \\ 0.17 \\ 0.30 \\ 0.39 \end{array} \right]$

$$\lambda_{\max} = 5.26; \text{C.I} = 0.07; \text{C.R} = 0.06$$

Further, VH and H effects demonstrated an intermediate ranking between weak and strong importance (average 3.96) when compared to a small effect; while a medium effect was considered of weak importance (2.90) over a small effect. Additionally, a VH and H effect were considered slightly less than of weak importance when compared to a medium effect (average 2.61). Finally, a VH effect demonstrated an intermediate ranking between equal and weak importance (1.90) when compared to a high effect.

Table 10.4 highlights pair-wise comparisons of levels of success with respect to the outcome social-emotional development. The table demonstrates that levels of success VH and H demonstrated an intermediate ranking between strong and very strong importance (average 5.87) when compared to no effect. A medium effect was considered of strong importance (4.7) compared to no effect. Moreover, a small effect demonstrated an intermediate ranking between weak and strong importance (3.8) on the outcome cognitive development.

**Table 10.4: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome social-emotional development (SED)**

<i>SED</i>	<b>N</b>	<b>S</b>	<b>M</b>	<b>H</b>	<b>VH</b>
<b>N</b>	1.00	0.26	0.21	0.18	0.17
<b>S</b>	3.80	1.00	0.30	0.28	0.25
<b>M</b>	4.70	3.30	1.00	0.42	0.39
<b>H</b>	5.70	3.63	2.40	1.00	0.47
<b>VH</b>	6.03	4.03	2.57	2.15	1.00

$\longrightarrow$ 

$\begin{bmatrix} \text{Priority} \\ 0.04 \\ 0.10 \\ 0.18 \\ 0.28 \\ 0.40 \end{bmatrix}$

$$\lambda_{\max} = 5.23; \text{C.I} = 0.06; \text{C.R} = 0.05$$

Additionally, a VH effect also demonstrated an intermediate ranking between weak and strong importance (4.03); while H and M effects are considered more than of weak importance (average 3.95) when compared to a small effect. Further, VH and H effects are considered slightly less than of weak importance (average 2.68) compared to a medium effect, and a VH effect demonstrated an intermediate ranking between equal importance and weak importance (2.15) compared to a high effect.

Table 10.5 highlights pair-wise comparisons of levels of success with respect to the outcome deviancy. Outcome levels of success VH, H, and M are considered slightly greater than of strong importance (average 5.55) compared to no effect, while a small effect is considered slightly less than a strong importance (4.5) compared to no effect. Moreover, a high effect is considered more important (5.70) when compared to no effect, in contrast to effects VH (5.45), M (5.5), and S (4.5).

**Table 10.5: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome deviancy (D)**

<i>D</i>	N	S	M	H	VH
N	1.00	0.22	0.18	0.18	0.18
S	4.50	1.00	0.30	0.25	0.25
M	5.50	3.30	1.00	0.31	0.29
H	5.70	4.00	3.20	1.00	0.36
VH	5.45	4.06	3.45	2.75	1.00

→

Priority
0.04
0.10
0.18
0.27
0.41

$$\lambda_{\max} = 5.56; \text{C.I} = 0.07; \text{C.R} = 0.06$$

Further, a very high and high effect demonstrated an intermediate ranking between weak and strong importance (average 4.03) compared to a small effect; while a medium effect is considered slightly greater than of weak importance (3.3) when compared to a small effect on the outcome deviancy. Additionally, very high and high effects were considered slightly greater than of weak importance compared to a medium effect (average 3.33). Finally, a very high effect is considered slightly less than of weak importance (2.75) compared to a high effect.

Table 10.6 highlights pair-wise comparisons of levels of success with respect to the outcome social participation. Outcome levels of success VH, H and M demonstrated an intermediate ranking between weak and strong importance (average 4.16) compared to no effect, while a small effect is considered less than of weak importance (2.4) compared to no effect. Outcome levels of success VH and H were only considered to be slightly greater than of weak importance (3.73) compared to a small effect, while a medium effect is considered slightly less than of weak importance (2.8) compared to a small effect.

**Table 10.6: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome social participation (SP)**

<i>SP</i>	<b>N</b>	<b>S</b>	<b>M</b>	<b>H</b>	<b>VH</b>
<b>N</b>	1.00	0.42	0.26	0.24	0.23
<b>S</b>	2.40	1.00	0.36	0.28	0.26
<b>M</b>	3.90	2.80	1.00	0.43	0.34
<b>H</b>	4.25	3.63	2.30	1.00	0.37
<b>VH</b>	4.33	3.83	2.90	2.70	1.00

$\longrightarrow$ 

Priority  
0.06  
0.09  
0.18  
0.25  
0.42

$$\lambda_{\max} = 5.26; \text{C.I} = 0.07; \text{C.R} = 0.06$$

Further, a very high effect is considered slightly less than of weak importance (2.90) compared to a medium effect with respect to the outcome social participation. A high effect demonstrated an intermediate ranking between equal importance and weak importance (2.3) when compared to a medium effect. Finally, a very high effect is considered slightly less than of weak importance compared to a high effect.

Table 10.7 compares outcome levels of success with respect to the domain criminal justice outcomes. Outcome levels of success VH and H were considered slightly greater than of strong importance (average 5.59) when compared to no effect. A medium effect is considered of strong importance (5.10) compared to no effect, while a small effect demonstrated an intermediate ranking between weak and strong importance (4.3) compared to no effect.

**Table 10.7: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome criminal justice outcomes (CJO)**

<i>CJ</i>	<b>N</b>	<b>S</b>	<b>M</b>	<b>H</b>	<b>VH</b>
<b>N</b>	1.00	0.23	0.20	0.18	0.17
<b>S</b>	4.30	1.00	0.32	0.27	0.23
<b>M</b>	5.10	3.10	1.00	0.26	0.26
<b>H</b>	5.45	3.73	3.80	1.00	0.29
<b>VH</b>	5.73	4.31	3.90	3.40	1.00

$\longrightarrow$ 

Priority  
0.04  
0.10  
0.16  
0.27  
0.43

$$\lambda_{\max} = 5.51; \text{C.I} = 0.13; \text{C.R} = 0.11$$

Further, very high and high effects demonstrated an intermediate ranking between weak and strong importance (4.02) compared to a small effect, while a medium effect is considered of weak importance when compared to a small effect. Additionally, a very high effect and a high effect (3.85) are considered slightly greater than of weak importance compared to a medium effect. Finally, a very high effect is considered only slightly greater than of weak importance (3.4) compared to a high effect.

Table 10.8 highlights pair-wise comparisons of levels of success with respect to the domain family wellbeing. Outcome levels of success VH and H are considered slightly greater than of strong importance (average 5.72) compared to no effect. Further, medium and small effects are considered slightly less than of strong importance (average 4.85) compared to no effect. A very high effect is seen as of weak importance compared to a small effect, while high and medium effects are considered slightly less than a weak importance (average 2.55) compared to a small effect on the outcome family wellbeing. Additionally, a high effect is considered more important than a very high effect (2.73 and

2.35 respectively) when compared to a medium effect. Finally, a very high effect demonstrated an intermediate ranking between equal importance and weak importance (1.93) compared to a high effect.

**Table 10.8: A matrix comparing outcome levels of success in pairs with respect to their perceived importance on the outcome family wellbeing (FW)**

<i>FW</i>	<b>N</b>	<b>S</b>	<b>M</b>	<b>H</b>	<b>VH</b>
<b>N</b>	1.00	0.22	0.20	0.18	0.17
<b>S</b>	4.60	1.00	0.41	0.37	0.31
<b>M</b>	5.10	2.42	1.00	0.37	0.43
<b>H</b>	5.50	2.68	2.73	1.00	0.52
<b>VH</b>	5.93	3.27	2.35	1.93	1.00

$\longrightarrow$ 

Priority  
0.04  
0.13  
0.17  
0.30  
0.36

$\lambda_{\max} = 5.24$ ; C.I = 0.06; C.R = 0.05

Figures 10.2 – 10.11 highlight preference values across all outcome domains with respect to pair-wise comparisons between the various outcome levels of success (e.g. a very high effect compared to no effect).

Figure 10.2 highlights the pair-wise comparison between the outcome levels of success labelled very high effect (VH) and no effect (N) across all outcome domains. Results indicate that the outcome domain educational success has the highest level of importance (6.4), which demonstrates an intermediate ranking between strong importance and very strong importance. The outcome domain social participation had the lowest level of importance (4.33), which generates an intermediate ranking between weak importance and strong importance.

**Figure 10.2: Pair-wise comparison between outcome levels of success very high effect (VH) and no effect (N)**

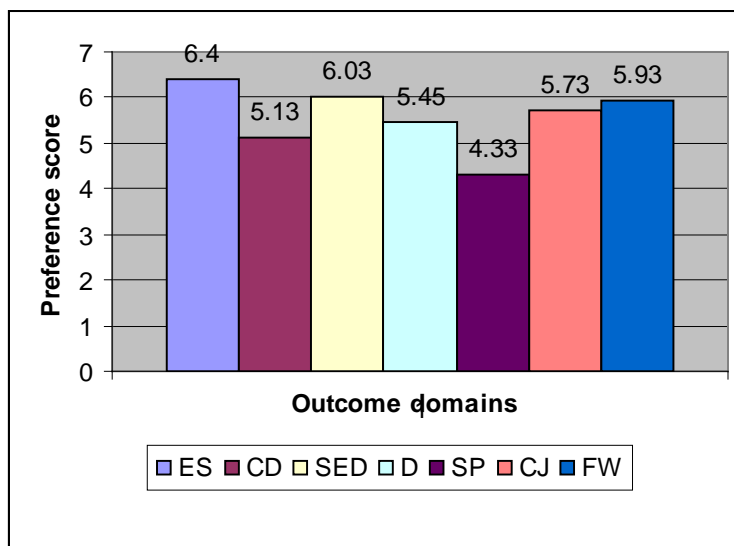


Figure 10.3 highlights the pair-wise comparisons between the outcome levels of success labelled high (H) and no effect (N) with respect to all outcomes. Results indicate that the outcome domain educational success has the highest level of importance (6.2), and in particular generates an intermediate ranking between a strong importance and a very strong importance. Further, the outcome domain social participation demonstrates the lowest level of importance (4.25), which produces an intermediate ranking between weak importance and strong importance on this outcome domain.



**Figure 10.3: Pair-wise comparison between outcome levels of success high effect (H) and no effect (N)**

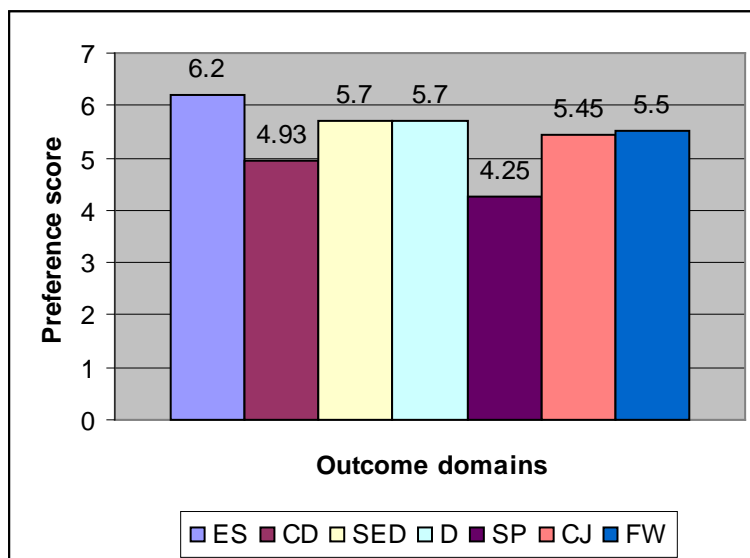


Figure 10.4 highlights the pair-wise comparisons between the outcome levels of success labelled M and N with respect to all outcome domains. Results indicate that the outcome domain deviancy is considered the highest level of importance (5.5), which indicates a strong importance over a medium effect compared to no effect on this domain. Further, the lowest preference score in this comparison is given to the outcome domain social participation (3.9), which produces an intermediate ranking between weak importance and strong importance.

**Figure 10.4: Pair-wise comparison between outcome levels of success medium effect (M) and no effect (N)**

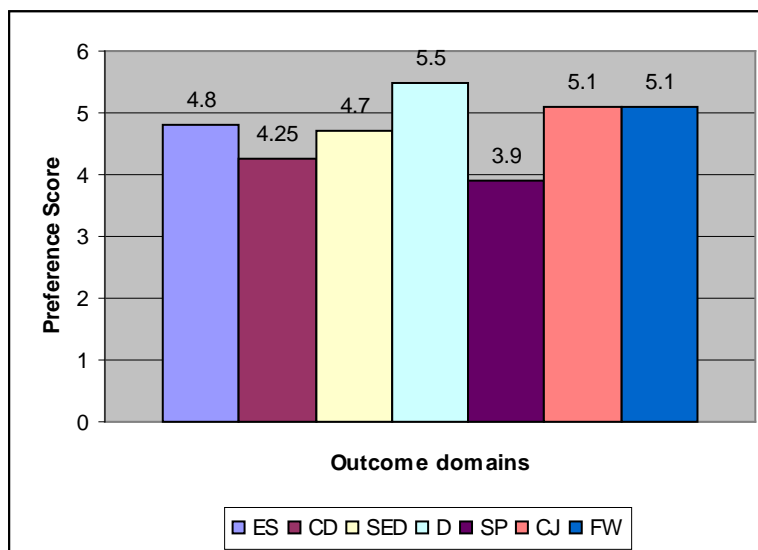


Figure 10.5 provides results of the pair-wise comparisons between the outcome levels of success labelled small (S) and no effect (N) across all outcome domains. Results highlight that the highest preference score across all outcome domains is given to the outcome family wellbeing (4.6) with a close second being the outcome deviancy (4.5). These results indicate that a small effect is considered slightly less than of strong importance on these outcomes compared to no effect. Further, the outcome domain social participation received the lowest rating of importance (2.4) with respect to the comparison between a small effect and no effect. This results highlights that a small effect demonstrates an intermediate ranking between equal importance and a weak importance when compared to no effect on this outcome domain.

**Figure 10.5: Pair-wise comparison between outcome levels of success small effect (S) and no effect (N)**

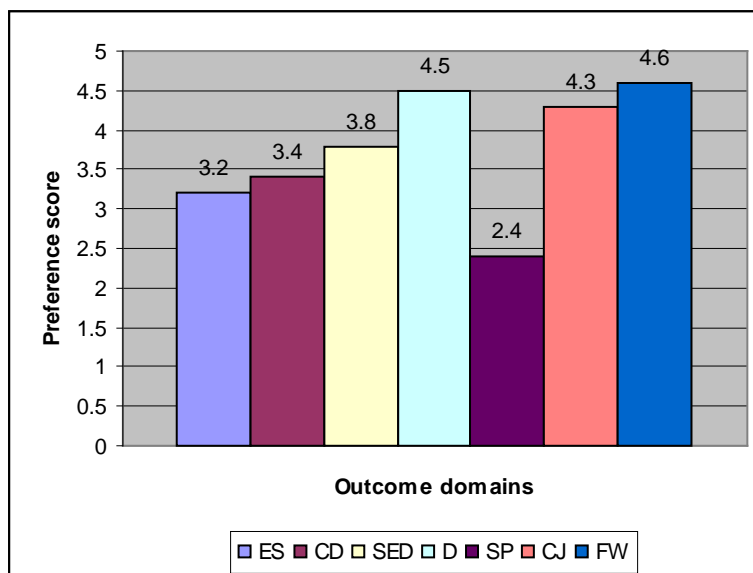


Figure 10.6 provides results of the pair-wise comparison between the outcome levels of success labelled very high (VH) and small (S) across all outcome domains. The highest score in this comparison is given to the outcome domain educational success. This indicates that a very high outcome is split between a weak importance and a strong importance (4.38) when compared with a small effect. Further, the outcome domain family wellbeing (FW) is given the lowest preference score (2.68), which highlights that the outcome FW is given an intermediate ranking between equal importance and weak importance when comparing a very high effect with a small effect on this outcome domain.

**Figure 10.6: Pair-wise comparison between outcome levels of success very high effect (VH) and small effect (S)**

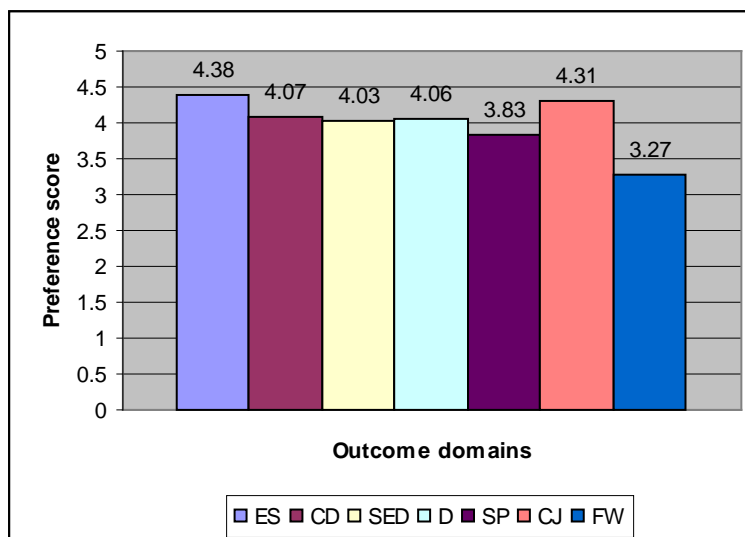


Figure 10.7 provides results of the pair-wise comparison between outcome levels of success high (H) and small (S) across all outcome domains. The outcome domain educational success is considered the most important with a preference value of (4.38), demonstrating an intermediate ranking between a weak importance and strong importance when comparing a high level of effect with a small effect on this outcome. Further, the outcome family wellbeing is given the lowest preference score (2.68) with respect to this pair-wise comparison. This result indicates a slightly less than weak importance when comparing a high effect with a small effect on this outcome domain.

**Figure 10.7: Pair-wise comparison between outcome levels of success high effect (H) and small effect (S)**

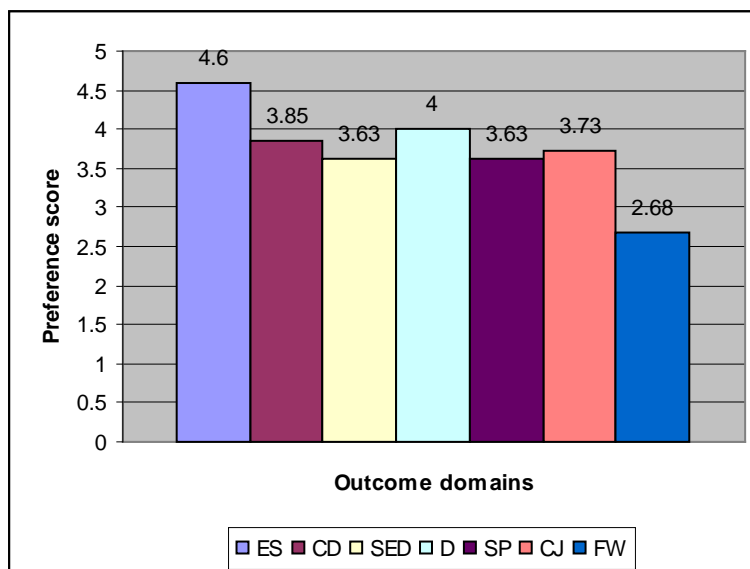


Figure 10.8 highlights that the pair-wise comparisons between the outcome levels of success labelled very high effect (M) and small effect (S) across all outcomes. Results indicate the outcome domains educational success, deviancy, and social-emotional development have the highest level of importance (3.3), which demonstrates a weak importance when comparing a medium effect with a small effect on these outcome domains. Moreover, the outcome domain family wellbeing received the lowest level of importance (2.42), which highlights that a medium effect is considered less than of weak importance over a small effect on this outcome.

**Figure 10.8: Pair-wise comparison between outcome levels of success medium effect (M) and small effect (S)**

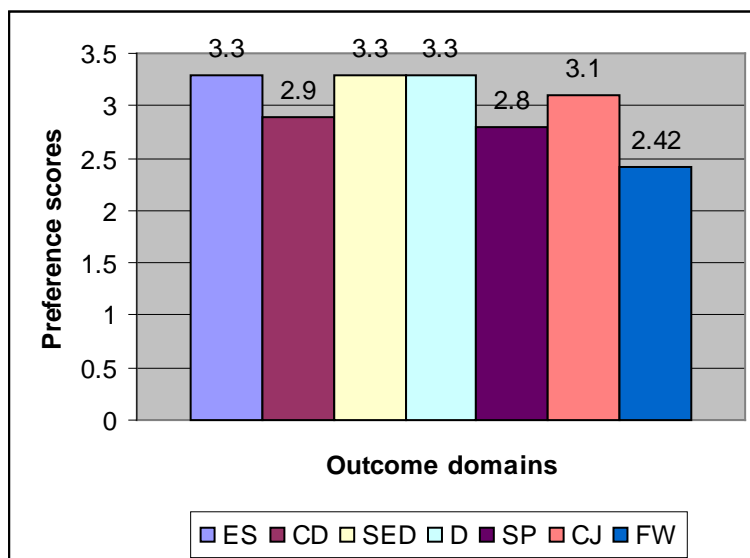


Figure 10.9 highlights the pair-wise comparisons between the outcome levels of success labelled VH and M with respect to all outcome domains. Results indicate that the outcome domain criminal justice is considered the highest level of importance (3.90) demonstrating an intermediate ranking between a weak importance and strong importance when comparing a very high level of effect with a medium effect on this outcome. Further, the lowest preference score in this comparison is given to the outcome domain family wellbeing (2.35), which corresponds to an intermediate ranking between equal importance and weak importance.

**Figure 10.9: Pair-wise comparison between outcome levels of success very high effect (VH) and medium effect (M)**

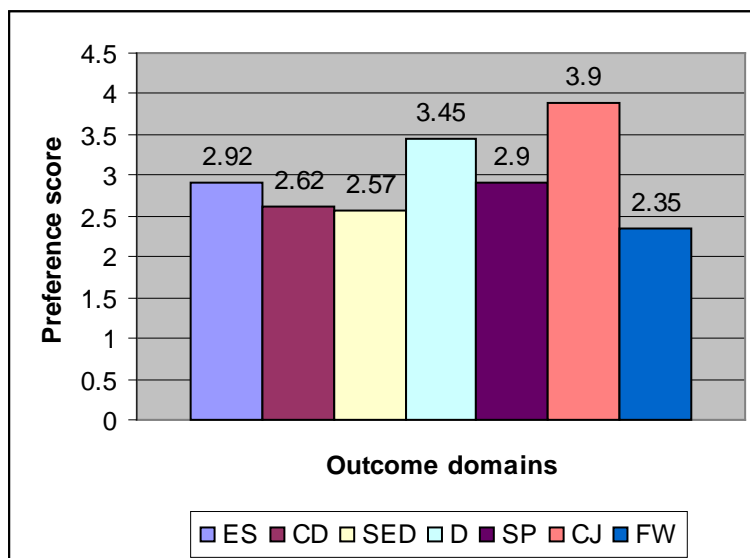


Figure 10.10 highlights the pair-wise comparisons between the outcome levels of success labelled H and M with respect to all outcome domains. Results indicate that the outcome domain criminal justice is considered the highest level of importance (3.8) demonstrating an intermediate ranking between a weak importance and strong importance when comparing a high level of effect with a medium effect on this outcome. Moreover, the outcome domain social participation received the lowest preference score (2.3) revealing an intermediate ranking between equal importance and weak importance when comparing a high level of effect with a medium effect on this outcome.

**Figure 10.10: Pair-wise comparison between outcome levels of success high effect (H) and medium effect (M)**

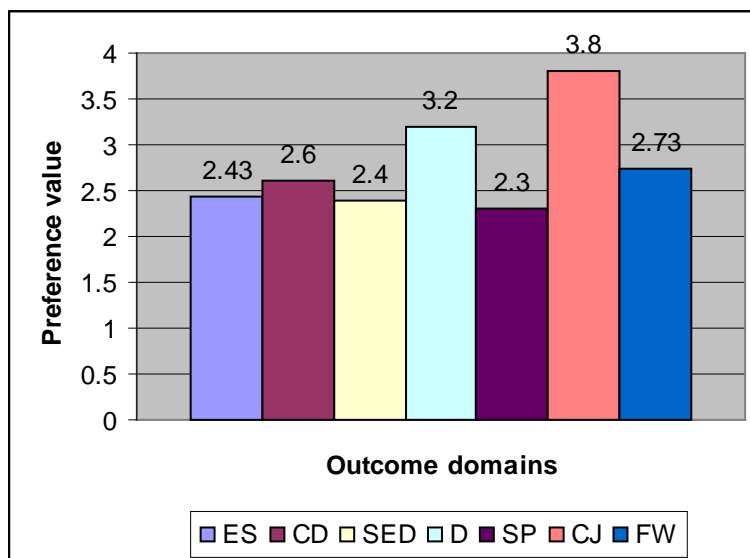
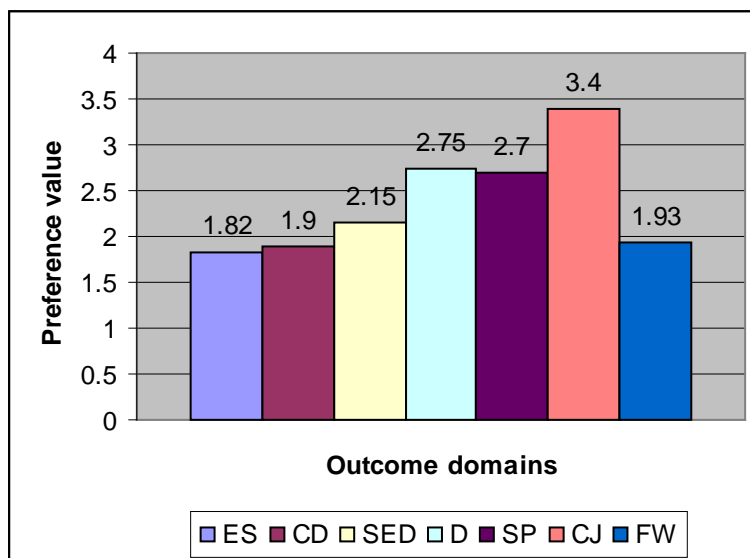


Figure 10.11 highlights the pair-wise comparisons between the outcome levels of success labelled VH and H with respect to all outcome domains. Criminal justice outcome are given the highest preference score (3.4) indicating a weak importance over a very high effect compared to a high effect on this outcome. Further, the outcome educational success received the lowest preference score (1.82) corresponding to an intermediate ranking between equal importance and weak importance when comparing a very high level of effect with a high effect on this outcome.



**Figure 10.11: Pair-wise comparison between outcome levels of success high effect (VH) and medium effect (H)**



Figures 10.2 to 10.11 demonstrate that the outcome domain educational success rated the highest preference score on sixty percent of occasions with respect to a higher level of success being compared with a lower level of success. Moreover, the outcome domain family wellbeing rated the lowest on forty percent of occasions with respect to a higher level of success being compared to a lower level of success. Additionally, the outcome domain deviancy scored the second highest preference score on sixty percent of comparisons when higher levels of success were compared with lower levels of success. Furthermore, the larger the gap between level of success (e.g. VH compared to S or N) the higher preference score it received. For example, when a very high effect was compared to no effect the average preference score was 5.57, indicating a strong importance. Conversely, a comparison between levels of success S and N and VH and H demonstrated only a weak importance (average 3.74) and an intermediate ranking between a equal importance and weak importance (average 2.38) respectively for all

outcomes. Interestingly, when reasonably large effects were being compared (e.g. VH and H, VH and M, and H and M) there was only a weak importance over one effect to the other across all outcome domains (average 2.71). Moreover, no pair-wise comparisons between outcome levels of success across all outcome domains received an absolute importance over another or a very strong importance over another. Consequently, this suggests that respondents believed that in most cases some effect was better than no effect. Consequently, when an effect of VH, H, M, and S were compared with no effect (5.57, 5.39, 4.76, and 3.74 respectively) levels of importance rated from of weak to strong importance.

A summary table combining the results from the seven matrices (Table 10.2 – 10.8) is provided (Table 10.9). Each column in Table 10.9 is weighted by the priority of the corresponding outcome domain (Section 10.2.1, Table 10.1) to obtain the weighted vector of priority for the intensity outcomes. A detailed example of the mechanics of achieving this weighting is provided in Section 8.4.

**Table 10.9: Summary table of the results comparing outcome levels of success with respect to their perceived importance on seven outcomes during adolescence**

	<b>ES (0.078)</b>	<b>CD (0.063)</b>	<b>SED (0.161)</b>	<b>D (0.113)</b>	<b>SP (0.144)</b>	<b>C (0.112)</b>	<b>FW (0.330)</b>
<b>N</b>	0.004	0.003	0.007	0.004	0.008	0.005	0.013
<b>S</b>	0.007	0.005	0.018	0.071	0.012	0.012	0.040
<b>M</b>	0.014	0.010	0.032	0.019	0.023	0.018	0.056
<b>H</b>	0.023	0.018	0.050	0.030	0.033	0.032	0.099
<b>VH</b>	0.032	0.023	0.072	0.045	0.055	0.052	0.119

Next, we selected from each column in Table 10.9 the element with the highest priority to obtain the vector of desired outcome levels with respect to their importance on each given outcome. We then divide each component by the sum of the vector. For example, normalised VH-ES =  $0.032/0.478 = 0.067$ . Table 10.10 provides normalised results for these priorities.

**Table 10.10: Vector of priorities for most preferred outcome/level of success**

<i>VH-ES</i>	0.032		<i>VH-ES</i>	0.067
<i>VH-CD</i>	0.023		<i>VH-CD</i>	0.048
<i>VH-SED</i>	0.072		<i>VH-SED</i>	0.151
<i>VH-D</i>	0.045	→	<i>VH-D</i>	0.094
<i>VH-SP</i>	0.055		<i>VH-SP</i>	0.115
<i>VH-CJ</i>	0.052		<i>VH-CJ</i>	0.109
<i>VH-FW</i>	0.199		<i>VH-FW</i>	0.416
<i>Total</i>	<b>0.478</b>		<i>Total</i>	<b>1.000</b>

Table 10.10 reveals that all respondents considered a very high effect most important when selecting an early intervention program based on its potential effect on a given outcome domain during the adolescent years. This is a particularly interesting result given that a very high effect (e.g. an effect size greater than 0.80, based on Cohen's (1992) description of effect sizes) is often difficult to achieve. Moreover, our meta-analysis revealed weighted mean effect sizes ranging from 0.148 - 0.532. Although some early intervention programs did demonstrate individual outcome effect sizes which could be considered high to very high, on average, effect sizes ranged from low to medium on non-health related outcomes during adolescence. Results from Table 10.10 highlight once again that family well-being (VH-FW) (0.416) followed by social-emotional

development (VH-SED) (0.151) is considered the highest priority with respect to their importance in contributing to NHRQOL during the adolescent years. The outcomes social participation (VH-SP) (0.115), (VH-CJ) criminal justice outcomes (0.109), (VH-D) deviancy (0.094), (VH-ES) educational success (0.067) and (VH-CD) cognitive development (0.048) constitute the remaining outcomes in order of preference with respect to the most desired outcome and intensity levels during the adolescent years. These results are consistent with those reported in Section 10.2.1.

### **10. 3: Results of survey 2**

Survey 2 determined the perceived preferences for early childhood intervention programs by developing seven matrices that compared the five preschool program options (structured preschool program (SPP), centre-based developmental day care (CBCC), home visitation (HV), family support services (FSS), and parental education (PE)) in pairs, with respect to the most desired attribute outcome/level of success (VH-ES, VH-CD, VH-SED, VH-D, VH-SP, VH-CJ, and VH-FW). Members of the school level group (e.g. school teachers and principals, co-ordinators of childcare centres, co-ordinators of behavioural management units, and co-ordinators of crèche and kindergartens) (n=8), and community agencies group (e.g. management and senior staff of private community organisations involved in the delivery of community-based developmental intervention programs - for example, Mission Australia) (n=7) were selected to complete the second survey. This survey comprised a total of 70 questions which took approximately 30 - 40 minutes for each respondent to complete. A copy of survey 2 is provided in Appendix J.

Results from the 15 respondents were aggregated and the average taken for each of seven outcomes. Tables 10.11–10.17 highlight the relative preference, strength and priorities among the five early childhood intervention programs (SPP, CBCC, HV, FSS, and PE) in pairs with respect to the most desired attribute outcome/level of success (VH-ES, VH-CD, VH-SED, VH-D, VH-SP, VH-CJ, and VH-FW). Appendix K provides descriptive statistics for all 70 questions.

Table 10.11 highlights that a structured preschool program (SPP) demonstrates an intermediate ranking between strong importance and very strong importance (5.6), and a weak importance (3.1) when compared with a centre-based childcare/developmental day care program (CBCC) and home visitation program (HV) respectively with respect to achieving a very high effect on the outcome educational success during adolescence. Moreover, a parental education program (PE) is considered of weak importance over CBCC with respect to achieving a very high effect on the outcome educational success. Further, SPP demonstrates an intermediate ranking between equal importance and weak importance when compared with family support (FSS) and parental education (PE) programs. Additionally, with respect to total priority rankings for this matrix, SPP is considered twice as important as PE, four times more important than CBCC, over three times more important than HV, and two and a half times more important than FSS. Finally, SPP and PE contribute to 62 percent of total priority in this matrix, compared to a combined total of 38 percent for CBCC, HV, and FSS for achieving a very high effect on the outcome educational success during adolescence.

**Table 10.11: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-ES.**

VH-ES	SPP	CBCC	HV	FSS	PE	
SPP	1.0	5.6	3.1	2.2	1.9	$\rightarrow \begin{bmatrix} \text{Priority} \\ 0.41 \\ 0.09 \\ 0.13 \\ 0.16 \\ 0.21 \end{bmatrix}$
CBCC	0.2	1.0	0.9	0.6	0.3	
HV	0.3	1.1	1.0	1.0	0.7	
FSS	0.4	1.7	1.0	1.0	0.9	
PE	0.5	3.2	1.5	1.1	1.0	

$$\lambda_{\max}=5.03; \text{C.I}=0.01; \text{C.R}=0.01$$

Table 10.12 indicates that SPP demonstrates an intermediate ranking between strong importance and very strong importance when compared with CBCC with respect to achieving a very high effect on the outcome cognitive development during adolescence. Further, SPP is considered of strong importance over HV, slightly greater than of weak importance over FSS and slightly less than of weak importance over PE with respect to achieving a very high effect on the outcome cognitive development. Moreover, CBCC is considered a weak importance over HV, and an intermediate ranking between equal importance and weak importance is demonstrated when CBCC is compared with FSS, FSS is compared with HV, PE is compared with CBCC, and PE is compared with FSS. Additionally, with respect to total priority rankings for this matrix, SPP constitutes approximately 50 percent of the total priority for this matrix (0.49%). CBCC (0.15%), HV (0.08%), FSS (0.11%), and PE (0.17%) make up the remainder of the priorities. Consequently, SPP is considered three times more important than CBCC, over six times more important than HV, over four times more important than FSS, and just less than three times more important than PE with respect to achieving a very high effect on the outcome cognitive development during adolescence.

**Table 10.12: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-CD.**

VH-CD	SPP	CBCC	HV	FSS	PE
SPP	1.0	6.1	4.9	3.7	2.5
CBCC	0.2	1.0	3.0	1.5	0.6
HV	0.2	0.3	1.0	0.9	0.5
FSS	0.3	0.7	1.2	1.0	0.8
PE	0.4	1.6	2.0	1.3	1.0

$\lambda_{\max}=5.18$ ; C.I = 0.04; C.R = 0.04

→

Priority  
 0.49  
 0.15  
 0.08  
 0.11  
 0.17

Table 10.13 highlights that SPP is considered slightly greater than of weak importance over CBCC (3.6) and HV (3.5) with respect to achieving a very high effect on the outcome domain social-emotional development during adolescence.

**Table 10.13: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-SED.**

VH-SED	SPP	CBCC	HV	FSS	PE
SPP	1.0	3.6	3.5	1.9	1.0
CBCC	0.3	1.0	1.7	1.1	0.8
HV	0.3	0.6	1.0	0.4	0.6
FSS	0.5	0.9	2.9	1.0	1.7
PE	1.0	1.2	1.6	0.6	1.0

$\lambda_{\max}=5.23$ ; C.I = 0.06; C.R = 0.05

→

Priority  
 0.35  
 0.15  
 0.09  
 0.22  
 0.19

This is a considerably lower preference than what was demonstrated in Tables 10.11 and 10.12 with respect to pair-wise comparisons between CBCC, HV and SPP. Further, SPP demonstrated an intermediate ranking between equal importance and weak

importance when compared with FSS (1.9) with respect to achieving a very high effect on the outcome social-emotional development. Moreover, intermediate rankings between equal importance and weak importance were seen when we compared the outcomes CBCC with HV (1.7), PE with HV (1.6), and FSS with PE (1.7). Additionally, PE and CBCC were seen as almost equally important to contributing to a very high effect on the outcome social-emotional development. With respect to total priority rankings for this matrix, SPP constitutes approximately 35 percent of the total priority for this matrix. This is considerably less than what was demonstrated in Tables 10.11 and 10.12. Further, FSS contributed to 22 percent of total priority. Combined, SPP, FSS and PE contributed to 76 percent of total priority. CBCC (0.15%), HV (0.09%) made up the remainder (0.24%). Consequently, SPP was considered over twice as important as CBCC and approximately four times more important than HV with respect to achieving a very high effect on the outcome social-emotional development during adolescence.

Table 10.14 highlights that SPP is considered slightly greater than a weak importance over CBCC (3.7) and HV (3.2) with respect to achieving a very high effect on the outcome domain deviancy during adolescence.



**Table 10.14: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-D.**

VH-D	SPP	CBCC	HV	FSS	PE	<div><div>Priority</div><div>0.30</div><div>0.10</div><div>0.12</div><div>0.25</div><div>0.23</div></div>
SPP	1.0	3.7	3.2	1.0	1.2	
CBCC	0.3	1.0	1.3	0.4	0.2	
HV	0.3	0.8	1.0	0.9	0.6	
FSS	1.0	2.2	1.1	1.0	2.0	
PE	0.8	4.3	1.8	0.5	1.0	
<div><div><math>\lambda_{\max}</math> =5.33; C.I = 0.08; C.R = 0.07</div></div>						

Once again, this is a considerably lower preference than what was demonstrated in Tables 10.11 and 10.12 with respect to pair-wise comparisons between CBCC, HV and SPP. Further, SPP and FSS are considered equally important to achieving a very high effect on the outcome deviancy. Moreover, a slightly greater than equal importance to contributing to a decrease in deviancy is given to comparisons between the programs SPP and PE (1.2), FSS and HV (1.1) and CBCC and HV (1.3). Additionally, parental education (PE) demonstrated an intermediate ranking between weak importance and strong importance when compared with CBCC (4.3); while comparisons between FSS and CBCC (2.2), and PE and HV (1.8) demonstrated an intermediate ranking between equal importance and weak importance with respect to achieving a very high effect on the outcome deviancy. With respect to total priority rankings for this matrix, early intervention programs SPP (0.30), FSS (0.25) and PE (0.23) contributed to 78 percent of total priorities; while the programs CBCC and HV contributed to only 22 percent of total priorities. Consequently, SPP was considered three times more important than CBCC and

slightly less than three times more important than HV with respect to achieving a very high effect on the outcome domain deviancy during the adolescent years.

Table 10.15 highlights that SPP is considered slightly less than a strong importance over CBCC (4.5) and slightly greater than a weak importance over HV (3.6) with respect to achieving a very high effect on the outcome social participation during adolescence.

**Table 10.15: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-SP.**

	SPP	CBCC	HV	FSS	PE
SPP	1.0	4.5	3.6	1.6	1.4
CBCC	0.2	1.0	1.9	1.0	0.5
HV	0.3	0.5	1.0	1.0	0.6
FSS	0.6	1.0	1.0	1.0	1.3
PE	0.7	2.0	1.6	0.8	1.0

$\longrightarrow$ 

Priority
0.36
0.14
0.12
0.18
0.20

$\lambda_{\max} = 5.24$ ; C.I = 0.06; C.R = 0.05

An intermediate ranking of importance between equal importance and weak importance was identified when we compared the early childhood programs SPP and FSS (1.6), CBCC and HV (1.9), PE and CBCC (2.0) and PE and HV (1.6) with respect to achieving a very high effect on the outcome social participation. Moreover, an equal importance or slightly greater than an equal importance was given to comparisons between SPP and PE (1.4), CBCC and FSS (1.0), FSS and HV (1.0), and FSS and PE (1.3) with respect to contributing to a very high effect on the outcome social participation. With respect to total priority rankings for this matrix, SPP contributed to 36

percent of total priorities; this was followed by PE (20%), FSS (18%), CBCC (14%), and HV (12%). Consequently, SPP was considered almost two times more important than PE, twice as important as FSS, two and a half times more important than CBCC and three times more important than HV with respect to achieving a very high effect on the outcome social participation during the adolescent years.

Table 10.16 highlights that SPP is considered slightly less than a strong importance over CBCC (4.6) and slightly less than a weak importance over HV (2.7) with respect to achieving a very high effect on the outcome criminal justice during the adolescent years. An intermediate ranking of importance between weak and strong importance was demonstrated when comparing PE with the early childhood intervention alternatives CBCC (3.7) and HV (4.2). Further, an intermediate ranking of importance between equal and weak importance was demonstrated when comparing FSS with the program alternatives CBCC (1.7), and HV (2.1). Additionally, equal importance was given to comparisons between the programs SPP and FSS (1.1), SPP and PE (1.2), and CBCC and HV (1.2) with respect to achieving a very high effect on criminal justice outcomes during adolescence. With respect to total priority rankings for this matrix, SPP, (0.30), FSS (0.26), and PE (0.26) contributed to 82 percent of total priorities. Moreover, CBCC and HV contributed 9 percent each to the remaining total. Consequently, SPP was considered over three times more important than CBCC and HV; and FSS and PE were considered just less than three times more important than CBCC and HV with respect to achieving a very high effect on criminal justice outcomes during adolescence.

**Table 10.16: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-CJO.**

	SPP	CBCC	HV	FSS	PE	Priority
SPP	1.0	4.6	2.7	1.1	1.2	0.30
CBCC	0.2	1.0	1.2	0.6	0.3	0.09
HV	0.4	0.9	1.0	0.5	0.2	0.09
FSS	0.9	1.7	2.1	1.0	1.9	0.26
PE	0.8	3.7	4.2	0.5	1.0	0.26

$\lambda_{\max}=5.23$ ; C.I = 0.06; C.R = 0.05

Table 10.17 highlights that SPP is considered slightly more than a weak importance and slightly less than a weak importance over the early intervention programs CBCC (3.3) and HV (2.5) with respect to achieving a very high effect on the family wellbeing outcome during adolescence. Further, an intermediate value between equal and weak importance was given to comparisons between the early intervention alternatives FSS and CBCC (1.9), FSS and HV (2.3), FSS and PE (1.9), and PE and CBCC (2.4) with respect to achieving a very high effect on this outcome domain.

**Table 10.17: A matrix comparing early childhood intervention programs in pairs with respect to the most desired attribute outcome/level of success for VH-FW.**

	SPP	CBCC	HV	FSS	PE	Priority
SPP	1.0	3.3	2.5	0.8	0.9	0.26
CBCC	0.3	1.0	1.2	0.5	0.4	0.11
HV	0.4	0.9	1.0	0.4	0.7	0.12
FSS	1.2	1.9	2.3	1.0	1.9	0.29
PE	1.1	2.4	1.5	0.5	1.0	0.22

$\lambda_{\max}=5.13$ ; C.I = 0.03; C.R = 0.03

Moreover, equal importance to achieving a very high effect on the outcome family wellbeing during adolescence was given to comparisons between the program alternatives CBCC and HV (1.2), FSS and SPP (1.2), PE and SPP (1.1), and PE and HV (1.5). With respect to total priority rankings for this matrix, FSS was considered to have the highest priority with respect to achieving a very high effect on the outcome domain family wellbeing (0.29). This was followed in order by SPP (0.26), PE (0.22), HV (0.12) and CBCC (0.11). Moreover, the program alternatives SPP, FSS and PE contributed to 77 percent of total priorities; with CBCC and HV contributing to only 23 percent overall. Consequently, FSS was considered to be two and a half times more important than the program alternatives HV and CBCC with respect to achieving a very high effect on the outcome family wellbeing during adolescence.

Results from Tables 10.11-10.17 demonstrate that structured preschool programs (SPP) tend to rate the highest priority with respect to their perceived importance to contributing to non-health related outcomes/intensities during the adolescent life phase (average priority 0.35). We noted that SPP programs rated the highest priority with respect to contributing to a very high effect on the outcome domain cognitive development (0.49) during adolescence. This stands in contrast to the outcome domains educational success, social-emotional development, deviancy, social participation, criminal justice outcomes and family wellbeing that had an average priority ranking of approximately 0.33. Notably, SPP rated 0.26 with respect to achieving a very high effect on the outcome domain family wellbeing during the adolescent years. FSS and PE ranked the next highest (average priority 0.21) with respect to potentially contributing to a very high effect on all outcomes during adolescence. Notably, FSS had fairly low priority

scores with respect to achieving a very high effect on the outcomes educational success (0.16), and cognitive development (0.11). The program alternatives CBCC and HV ranked, on average, the lowest on priority rankings across the seven outcomes. However, the program alternative CBCC ranked third overall (0.15) with respect to achieving a very high effect on the outcome cognitive development during adolescence.

Table 10.18 summarises the results from the seven matrices given in Tables 10.11 to 10.17. It is derived by multiplying the priority vectors from Tables 10.11 to 10.17 by the corresponding vector of priorities for most preferred outcome/level of success (given in Table 10.10); thus obtaining the weighted vector of priority for all program alternatives with respect to the various intensity/outcomes (e.g. SPP with respect to VH-ES, VH-CD, VH-SED, VH-D, VH-SP, VH-CJ, VH-FW and so on). For example, the priority score 0.41 (given in Table 10.11, column seven, row one) is multiplied by the vector of priority VH-ES (0.067) (given in Table 10.10), which results in a weighted priority for the program alternative SPP with respect to the outcome intensity VH-ES (0.03) as recorded in row 1, column 1 of Table 10.18. A detailed example of the mechanics of deriving these weighted priorities is provided in Section 8.4.

**Table 10.18: Weighted priorities for the early childhood intervention program options with respect to outcome/intensities**

	VH-ES	VH-CD	VH-SED	VH-D	VH-SP	VH-CJ	VH-FW
<b>SPP</b>	0.03	0.02	0.05	0.03	0.04	0.03	0.11
<b>CBCC</b>	0.01	0.01	0.02	0.01	0.02	0.01	0.05
<b>HV</b>	0.01	0.00	0.01	0.01	0.01	0.01	0.05
<b>FSS</b>	0.01	0.01	0.03	0.02	0.02	0.03	0.12
<b>PE</b>	0.01	0.01	0.03	0.02	0.02	0.03	0.09

Moreover, we added each of the weighted elements in each row to obtain the overall priorities of the five early childhood intervention program alternatives (e.g. SPP\*, CBCC\*, HV\*, FSS\*, and PE\*) ( see Equation 10.1).

$$\begin{aligned} SPP^* = & SPP/VH-ES (0.03) + SPP/VH-CD (0.02) + SPP/VH-SED (0.05) + SPP/VH-D \\ & (0.03) + SPP/VH-SP (0.04) + SPP/VH-CJ (0.03) + SPP/VH-FW (0.11) \end{aligned} \quad (10.1)$$

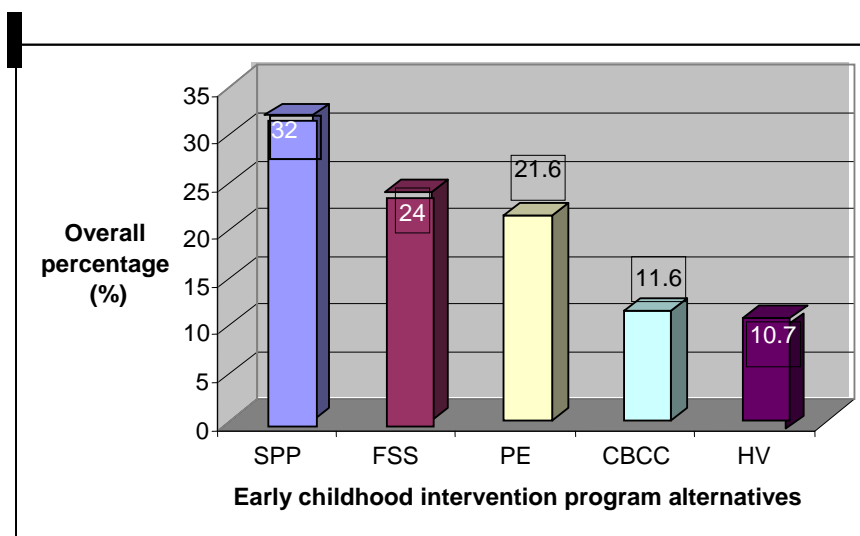
The synthesis of each row yields the following priorities in order of highest priority to lowest priority: SPP\* (0.320), FSS\* (0.240), PE\* (0.216), CBCC\* (0.116), and HV\* (0.107) (see Table 10.19)

**Table 10.19: Overall priorities of five early childhood intervention program alternatives with respect to very high effect on seven outcome domains**

	<i>Priority</i>
<i>SPP *</i>	0.320
<i>FSS *</i>	0.240
<i>PE *</i>	0.216
<i>CBCC *</i>	0.116
<i>HV *</i>	0.107

Table 10.19 and Figure 10.12 highlight that a structured preschool program (SPP) is considered the highest priority with respect to contributing to a very high effect on all outcomes (ES, CD, SED, D, SP, CJ, and FW) during the adolescent years. This is followed in order of priority by the early intervention programs family support services (FSS) (0.240), parental education (PE) (0.216), centre-based childcare/ developmental day care (CBCC) (0.116), and home visitation (HV) (0.107).

**Figure 10.12: Percentage of priorities of five early childhood intervention program alternatives with respect to a very high effect on seven outcome domains**



Consequently, SPP is considered three times more important than HV, and just less than three times more important than CBCC with respect to achieving a very high effect on all outcomes during the adolescent years. FSS and PE programs are considered near equal with respect to their potential contribution to achieving a very high effect on all outcomes during adolescence. In turn FSS and PE programs are seen as more than twice as important as the program alternatives CBCC and HV with regard to achieving a very high effect on all outcomes during the adolescent years.

#### **10.4: Synthesis of surveys 1 and 2 highlighting priorities for five early childhood intervention programs**

Figure 10.13 provides a schematic representation of the final results reported in Section 10.3. Overall, structured preschool programs (SPP) (0.320) were considered the highest priority with respect to their potential positive contribution to non health-related



quality of life during adolescence. SPP was followed in order of preference by family support services (FSS) (0.240), parental education (PE) (0.216), centre-based childcare (0.116) and home visitation (HV) (0.107).

**Figure 10.13: Synthesis of priorities**



The overall level of inconsistency for the hierarchy is 0.06 which is considered acceptable, highlighting that choices made by respondents were overall relatively consistent. We obtained this figure using the method outlined in Section 8.2.3. Further, we checked this value using the software ExpertChoice (Expert Choice, 2000-2004). This value is important as it suggests that respondents generated transitive social ordering of preferences. A transitive ordering of preferences, as outlined in Chapter 9, requires that if  $X$  is preferred to  $Y$ , and  $Y$  is preferred to  $Z$ , then  $X$  is preferred to  $Z$ . This logic of transitivity is taken as an axiom of rationality with regards to preferences of individuals (Boardman et al., 1996, 2001, 2006). A violation of transitivity implies an ambiguous ordering of alternatives. Moreover, based on the axioms of the analytic hierarchy process (Saaty, 1980, 1986) a consistency of 0.10 or less is considered acceptable.

When results are normalised (Figure 10.14), we can see that if a structured preschool program (SPP) is given a value of 1, it is almost three times as preferred as home visitation (HV) and centre-based child care/developmental day care (CBCC).

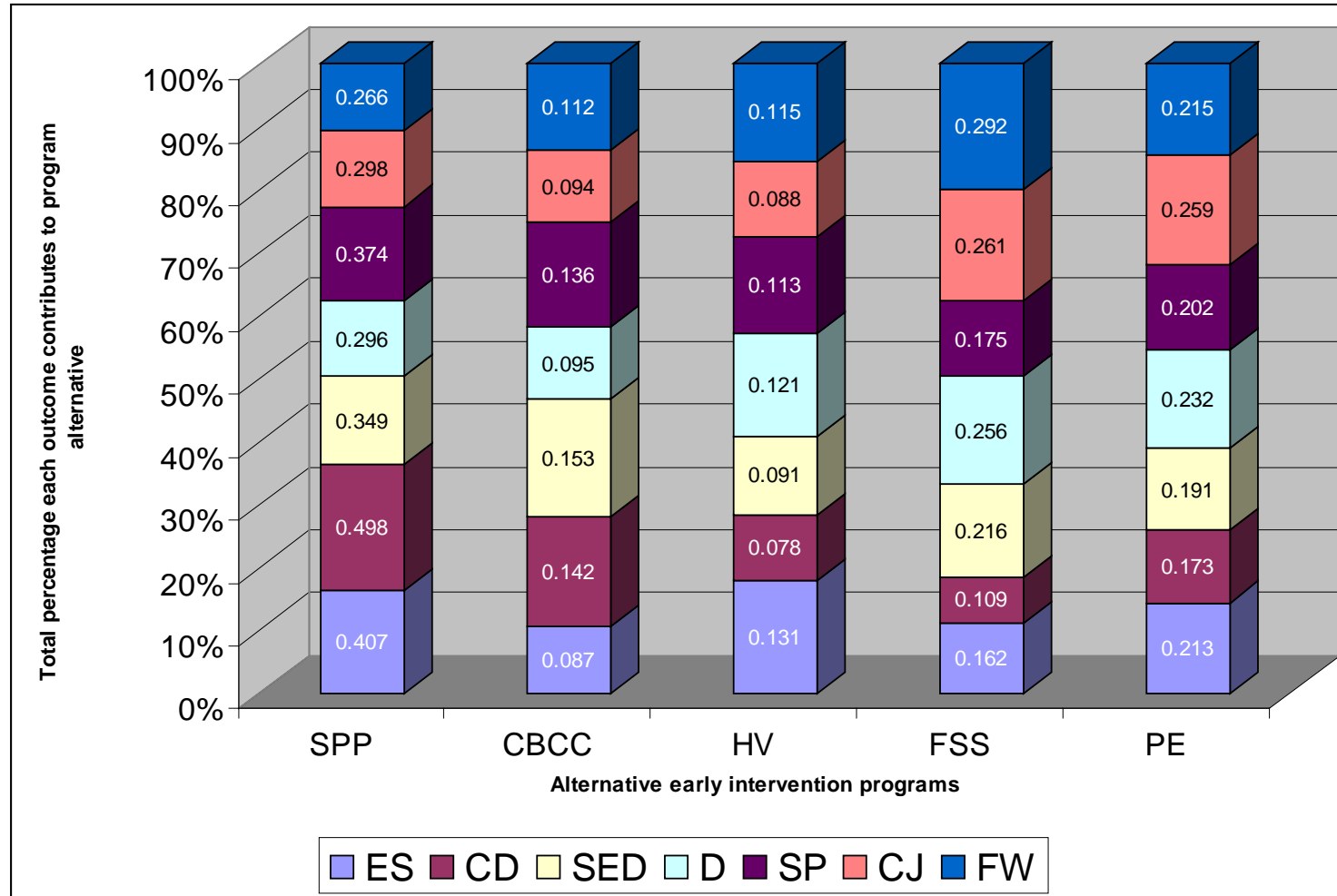
Moreover, parental education (PE) and family support services (FSS) are more than twice as preferred as centre-based childcare/developmental day care (CBCC) and home visitation (HV) programs.

**Figure 10.14: Normalising of the synthesised priorities**



Figure 10.15 provides a schematic representation of the hierarchy incorporating results from level 1 of the hierarchy (pair-wise comparisons between outcomes during adolescence and their potential contribution to non health-related quality of life) and their overall contribution, in terms of their percentage of priority, to each early childhood intervention program alternative. With respect to the program alternative SPP (overall priority .320) (see Figure 10.13), the outcome cognitive development (CD) contributed the highest priority (0.498), followed in order of importance by the outcomes educational success (0.407), social participation (0.374), social-emotional development (0.349), criminal justice outcomes (0.298) and family wellbeing (0.266). In regards to the program alternative CBCC (overall priority 0.116) (see Figure 10.13), the outcome social-emotional development (SED) contributed the highest priority (0.153), followed in order of importance by the outcomes cognitive development (0.142), social participation (0.136), family wellbeing (0.112), deviancy (0.095), criminal justice outcomes (0.094), and educational success (0.087).

**Figure 10.15: Summary of the hierarchy incorporating results from level 1**



With respect to the program alternative HV (overall priority 0.107) (see Figure 10.13), the outcome educational success (ES) contributed the highest priority (0.131), followed in order of importance by the outcomes deviancy (0.121), family wellbeing (0.115), social participation (0.113), social-emotional development (0.091), criminal justice outcomes (0.088), and cognitive development (0.078). In regards to the program alternative FSS (overall priority 0.240) (see Figure 10.13), the outcome family wellbeing (FW) contributed the highest priority (0.292), followed in order of importance by the outcomes criminal justice outcomes (0.261), deviancy (0.256), social-emotional development (0.216), social participation (0.175), educational success (0.162), and cognitive development (0.109).

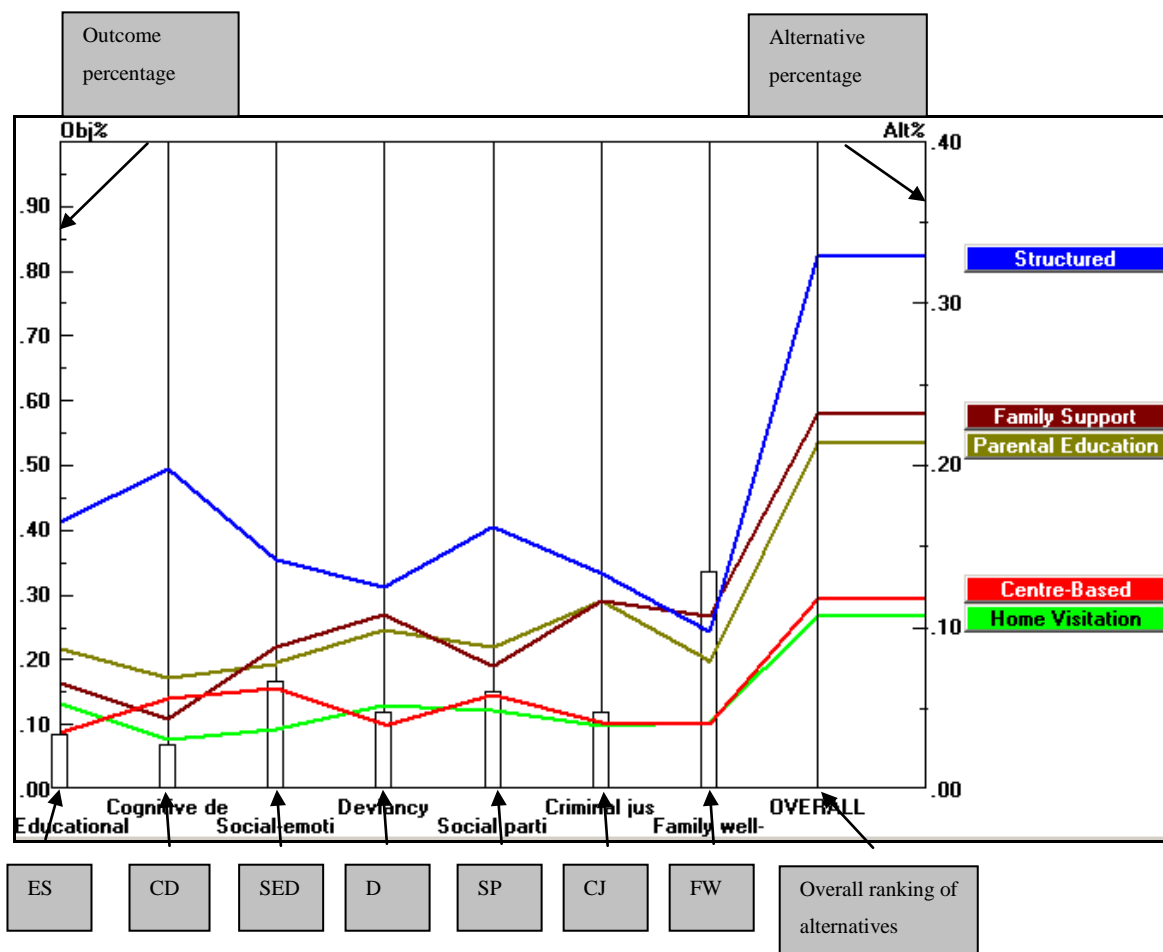
Finally, with respect to the program alternative parental education (overall priority 0.216) (see Figure 10.13), the outcome criminal justice outcomes (CJ) contributed the highest priority (0.259), followed in order of importance by the outcomes deviancy (0.232), family wellbeing (0.215), educational success (0.213), social participation (0.202), social-emotional development (0.191), and cognitive development (0.173).

Figure 10.16 combines results of the priority rankings from surveys 1 and 2. The vector of priorities derived from outcome objectives (ES, CD, SED, D, SP, CJ, and FW) are provided on the *X*-axis (schematically represented by vertical bars). The overall outcome percentage is provided left hand *Y*-axis, which ranges from .00 to .90. Further, program alternatives (SPP, CBCC, HV, FS, and PE) are provided on the right hand *Y*-axis, together with the alternative priority percentage, which ranges from .00 to .40. Moreover, overall rankings of program alternatives can be read from the right

hand Y-axis. The coloured lines represent the overall contribution made by non health-related outcomes, in terms of their percentage of priority, to each early childhood intervention program alternative.

Figure 10.16 demonstrates that early intervention program SPP rates the highest (32%) with respect to overall priority percentage rankings and its potential contribution to improving non health-related quality of life during adolescence. This is followed in order of percentage priority (highest to lowest) by the program alternatives FSS (24%), PE (21.6%), CBCC (11.6%), and HV (10.7%).

**Figure 10.16: Overall performance graph for hierarchy**



Moreover, Figure 10.16 highlights that the non health-related outcome FW is rated the highest with respect to its potential contribution to improving non health-related quality of life during adolescence (33% of total priority); followed in order of percentage priority by the non health-related outcomes SED (16.1%), SP (14.4%), D (11.3%), CJ (11.2%), ES (7.8%), and CD (6.3%). Furthermore, Figure 10.16 allows us to identify what outcomes contributed the most to the percentage priority rankings of the five early childhood intervention program alternatives. This is represented by the coloured lines showing the overall contribution made by non health-related outcomes, in terms of their percentage of priority, to each program alternative. The coloured lines demonstrate a large gap between the outcomes ES, CD, SED and SP, whereby the blue line (representing structured preschool programs) is separated from the red, brown, khaki, and green lines. Consequently, this shows us that SPP programs received large priority percentage ratings with respect to the non health-related outcomes ES, CD, SED and SP. In contrast, we can see that the red (CBCC program) and green (HV program) lines separate from the other coloured lines with regards to the non health-related outcomes deviancy (D) and criminal justice outcomes (CJ). Thus, the outcomes D and CJ received rather low priority percentage rankings with respect to their overall contribution to the program alternatives CBCC and HV. Consequently, these outcomes influenced the overall percentage of priority rankings between CBCC and HV with respect to the program alternatives SPP, FSS, and PE.

### **10.5: Sensitivity analysis**

The stability of the results was tested using a simple sensitivity analysis. The sensitivity analysis looks at ‘what if’ scenarios and allows one to measure the responsiveness of the results to changes in the relative importance of one or more of

the outcome objectives. A change in the relative importance of one of the outcome objectives will create changes to the relative importance of the overall alternatives with respect to the goal. In this thesis, this goal is the perceived contribution to non-health related quality of life during adolescence. The sensitivity analysis will produce effects on the hierarchy as a whole and may create a new set of priorities with respect to the overall importance of each of the early childhood intervention alternatives.

Figure 10.16 demonstrates that the outcome objective family well-being (FW) has the highest priority, followed in order of priority (highest to lowest) by the outcomes social-emotional development (SED), social participation (SP), criminal justice outcomes, educational success (ES), and cognitive development (CD) respectively. With the outcome objectives weighted this way, structured preschool programs (SPP) was the most preferred option among the program alternatives. In our first ‘what if’ scenario, we reduce the most preferred outcome objective (FW) by an overall ten percentage points to see how a change in the relative importance of one of the outcome objectives (in this case FW) will create changes to the relative importance of the overall program alternatives with respect to the goal. Additionally, we employ a gradient sensitivity analysis to identify the cross-over points of the program alternatives, which indicates the direction and strength of a priority change required to modify the percentage priority rankings of the five early childhood intervention program alternatives.

Figure 10.17 demonstrates that if we reduce the most preferred outcome objective (FW) by ten percent there is no change to the ranking of program alternatives. In this case, the SPP program increases by an overall 0.08 percent of overall ranking priorities. Centre-based programs increase by 0.01 percent, home visitation remains stable at 10.7 percent, family support services decline by 0.08 percent and parental education programs remain stable at 21.6 percent.

**Figure 10.17: Sensitivity for early childhood intervention alternatives (Reducing the most preferred outcome (FW) by 10%)**

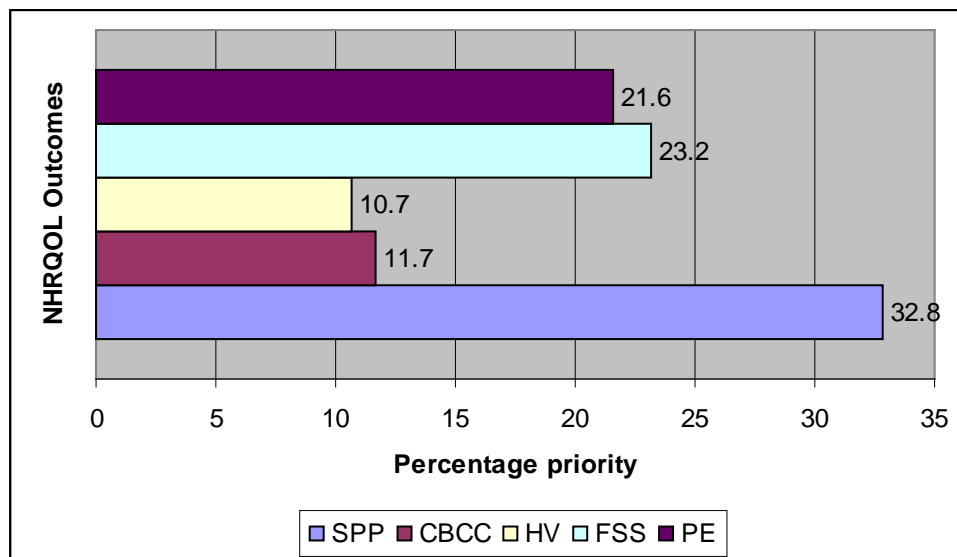




Figure 10.18 provides a schematic representation of the overall effects on outcomes as a result of our first ‘what if’ scenario. In this case, results demonstrate that although overall performance on the outcome domain FW has reduced, there has been no change to the ranking of early intervention program alternatives overall and little change with respect to each outcome objective.

**Figure 10.18: Overall performance graph for hierarchy with a 10 percent reduction in the outcome FW**

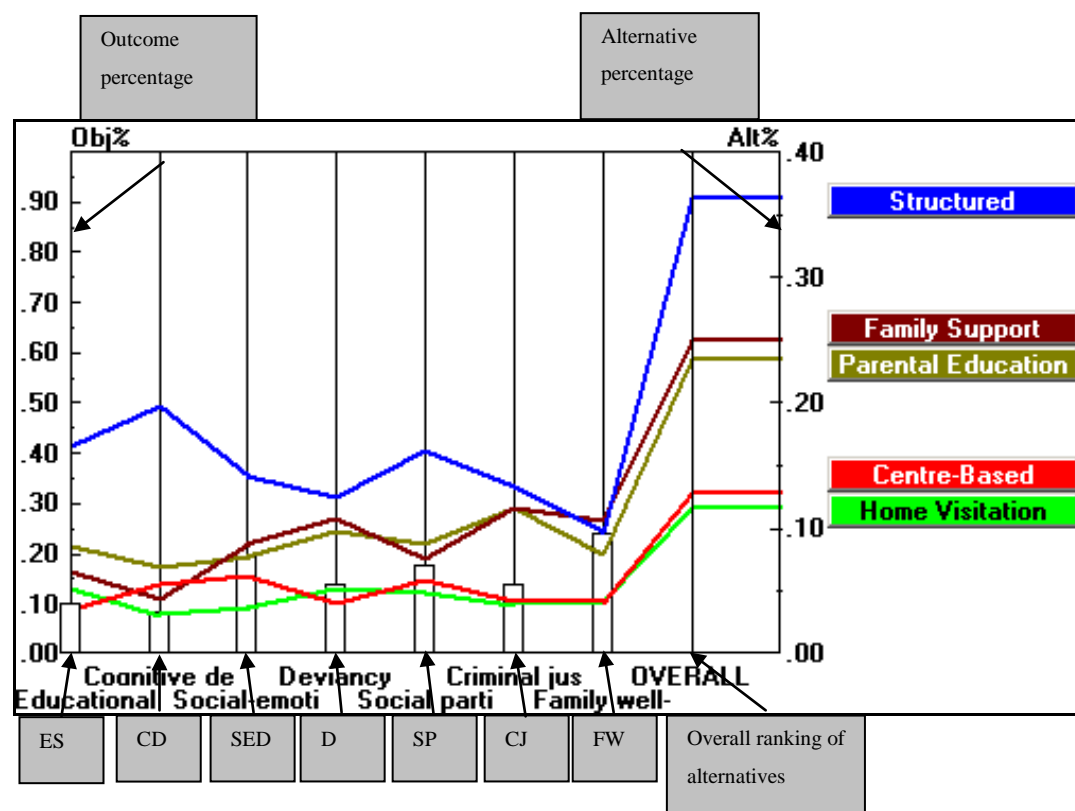
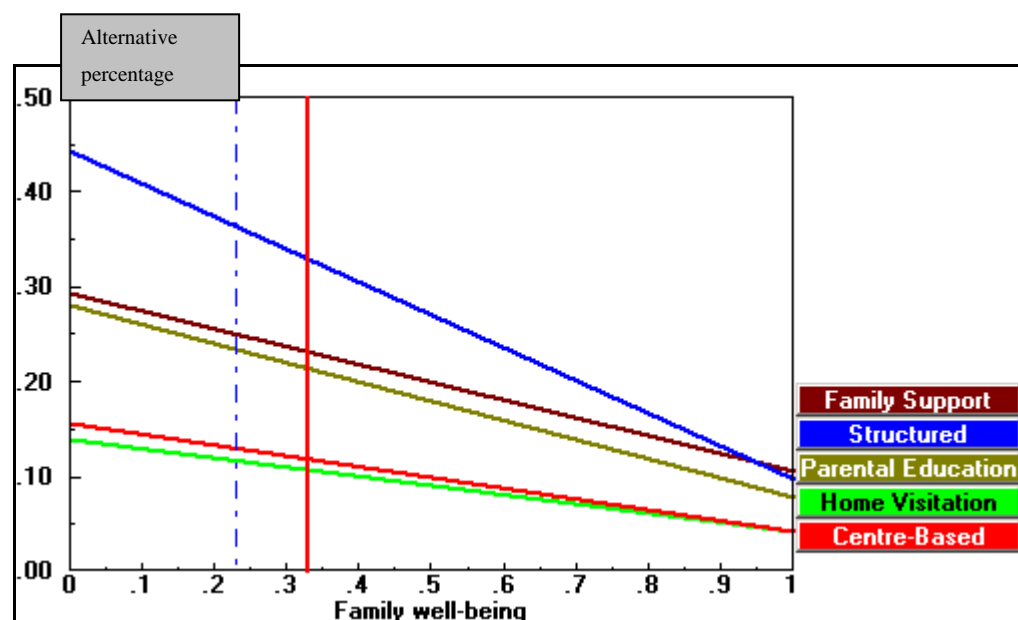


Figure 10.19 shows the alternative program priorities (SPP, CBCC, HV, FSS and PE) with respect to a ten percent reduction in the priority rating of the outcome objective FW. The red vertical line indicates the outcome objectives priority (33%) based on the decision-maker's paired comparisons. The dotted line demonstrates the change in the outcome objectives priority (23.1%).

**Figure 10.19: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome FW and possible changes in program priority rankings**



In this figure, we are looking for the cross-over points of the program alternatives, which indicates the direction and strength of a priority change (in this case in the outcome FW) required to change the percentage priority rankings of the five early childhood intervention program alternatives. The figure demonstrates that we are a significant distance from any of these cross-over points whereby the ranking of program alternatives would change. Moreover, the outcome FW would have to have a percentage priority ranking of approximately 91.5% for a structured preschool

program to be rated lower than a family support program with respect to its potential contribution to non health-related quality of life during adolescence.

Figures 10.20 to 10.25 provide gradient sensitivity graphs for our first sensitivity analysis, namely the effect of a 10 percentage point reduction in FW on the six remaining outcomes CJ, ES, CD, SED, D, and SP outcomes and overall changes in program alternative rankings.

Figure 10.20 demonstrates that changes in percentage of priority ranking on the outcome domain criminal justice outcomes will not change the overall priority rankings of the five early childhood intervention programs alternatives. Further, the figure shows that a reduction on the overall percentage priority rating of FW (10 % in sensitivity scenario 1) has an effect on the outcome CJ, thereby increasing the overall percentage priority of criminal justice outcomes to 12.9 percent (an increase of 1.7%).

**Figure 10.20: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome CJO and possible changes in program priority rankings**

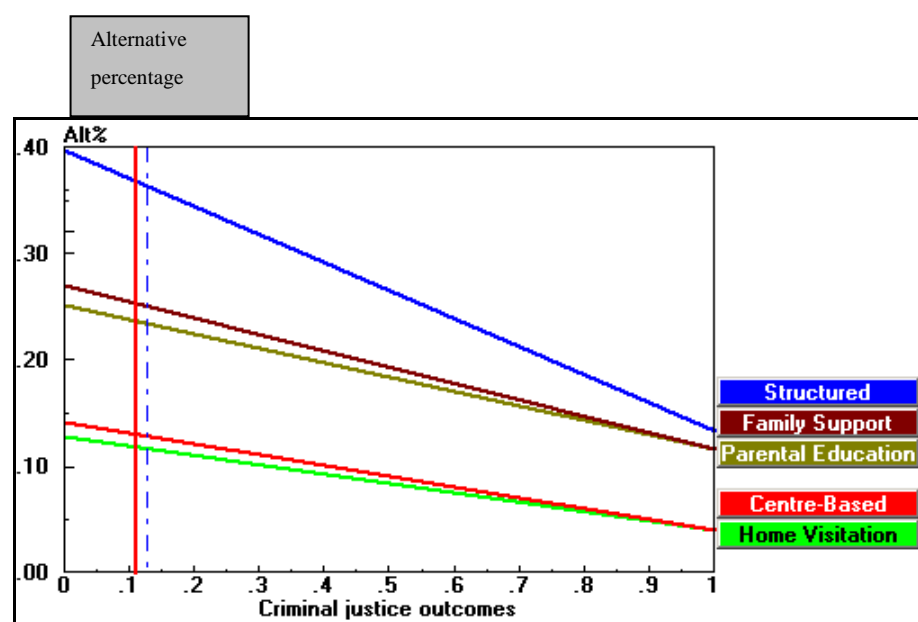


Figure 10.21 demonstrates that an increase in overall percentage priority outcome educational success will have the potential to change overall priority rankings on the program alternatives. In particular the figure shows that a reduction on the overall percentage priority rating of FW (10% in sensitivity scenario 1) has an effect on the outcome educational success, thereby increasing the overall percentage priority of educational success to 9 percent (an increase of 1.2%). However, Figure 10.21 also highlights that the outcome educational success would have to increase its overall percentage priority ranking by approximately 30 percent for any change to occur in the ranking of program alternative rankings.

**Figure 10.21: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome ES and possible changes in program priority rankings**

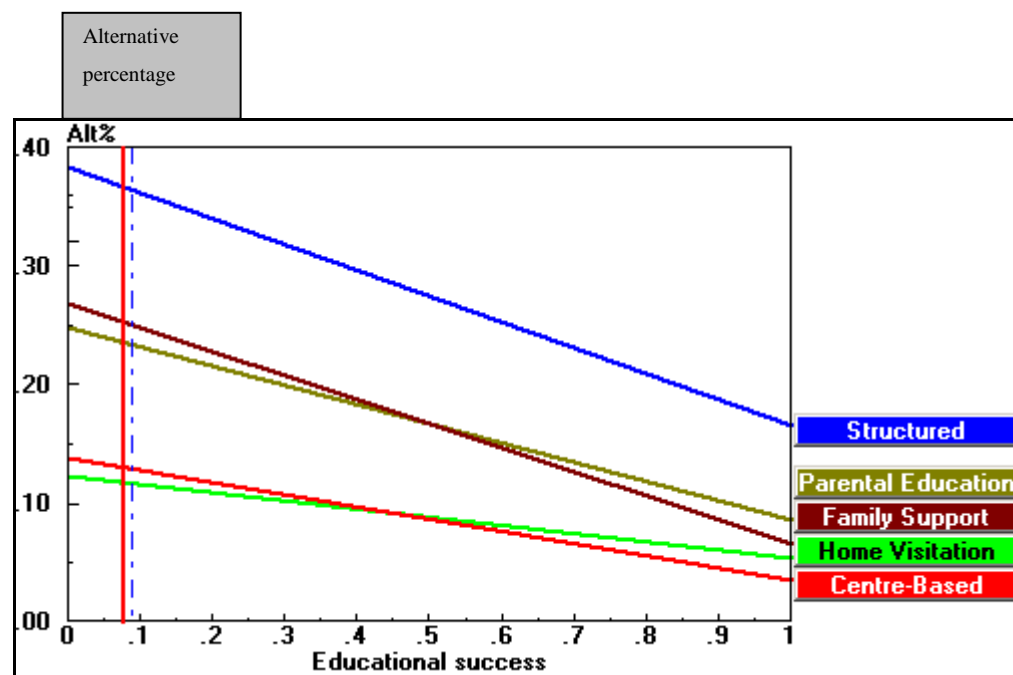


Figure 10.22 highlights that an increase in overall percentage priority outcome cognitive development will have the potential to change overall priority rankings on the program alternatives. In particular this figure shows that a reduction on the overall percentage priority rating of FW (10% in sensitivity scenario 1) has an effect on the outcome cognitive development, thus increasing the overall percentage priority of cognitive development to 7.2 percent (an increase of 0.9%). Furthermore, Figure 10.22 also demonstrates that the outcome cognitive development would have to increase its overall percentage priority ranking by approximately 25 percent for any change to occur in the ranking of program alternatives family support and parental education.

**Figure 10.22: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome CD and possible changes in program priority rankings**

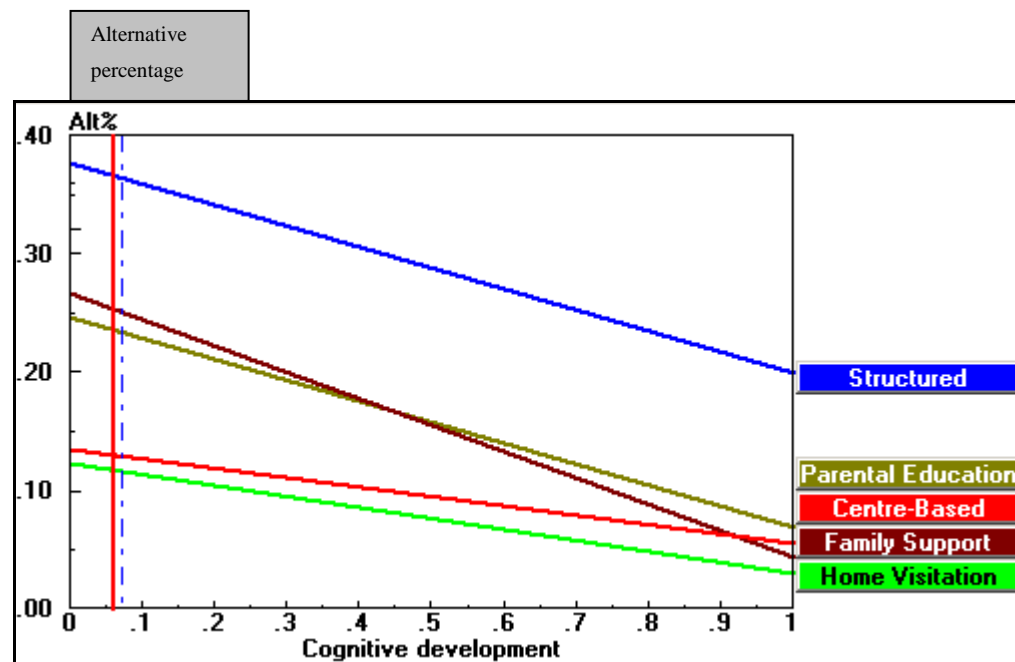


Figure 10.23 demonstrates that changes in percentage of priority ranking on the outcome domain social-emotional development (SED) will not change the overall priority rankings of the five early childhood intervention programs alternatives. However, the figure does reveal that a reduction on the overall percentage priority rating of FW (10% in sensitivity scenario 1) has an effect on the outcome SED. In particular it generates an increase in the overall percentage priority of social-emotional development outcomes to 18.5 percent (an increase of 2.4%).

**Figure 10.23: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome SED and possible changes in program priority rankings**

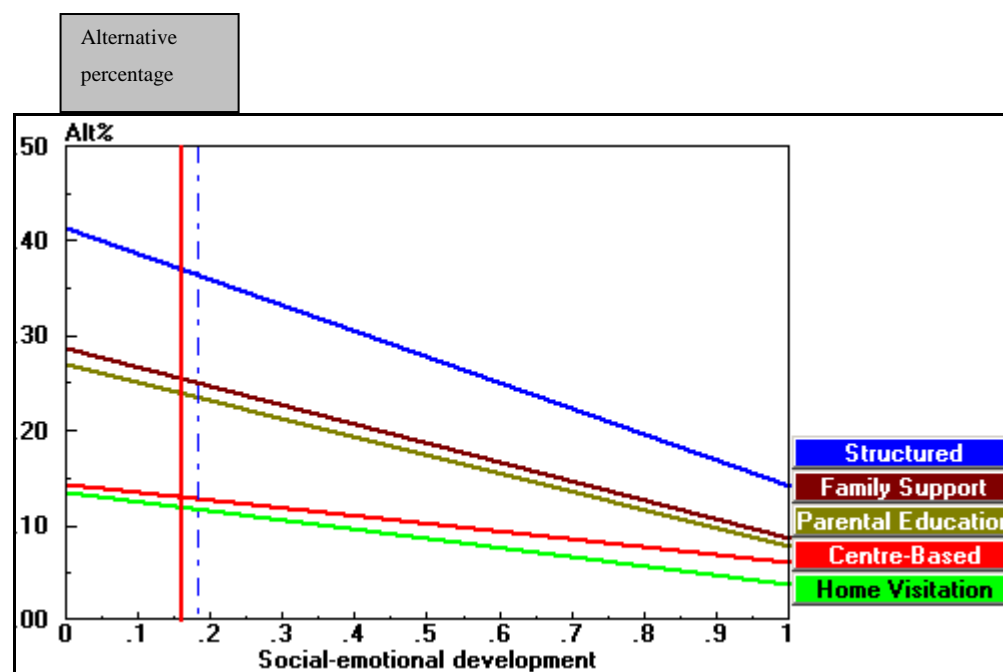


Figure 10.24 highlights that an increase in overall percentage priority outcome deviancy will have the potential to change overall priority rankings of the program alternatives. The figure shows that a reduction in the overall percentage priority rating of FW (10% in sensitivity scenario 1) has an effect on the outcome deviancy. In particular it generates an increase in the overall percentage priority of deviancy to 12.9 percent (an increase of 1.6%). Furthermore, Figure 10.24 demonstrates that the outcome deviancy would have to increase its overall percentage priority ranking by approximately 35 percent for any change to occur in the ranking of program alternatives home visitation and centre-base childcare/developmental day care.

**Figure 10.24: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome D and possible changes in program priority rankings**

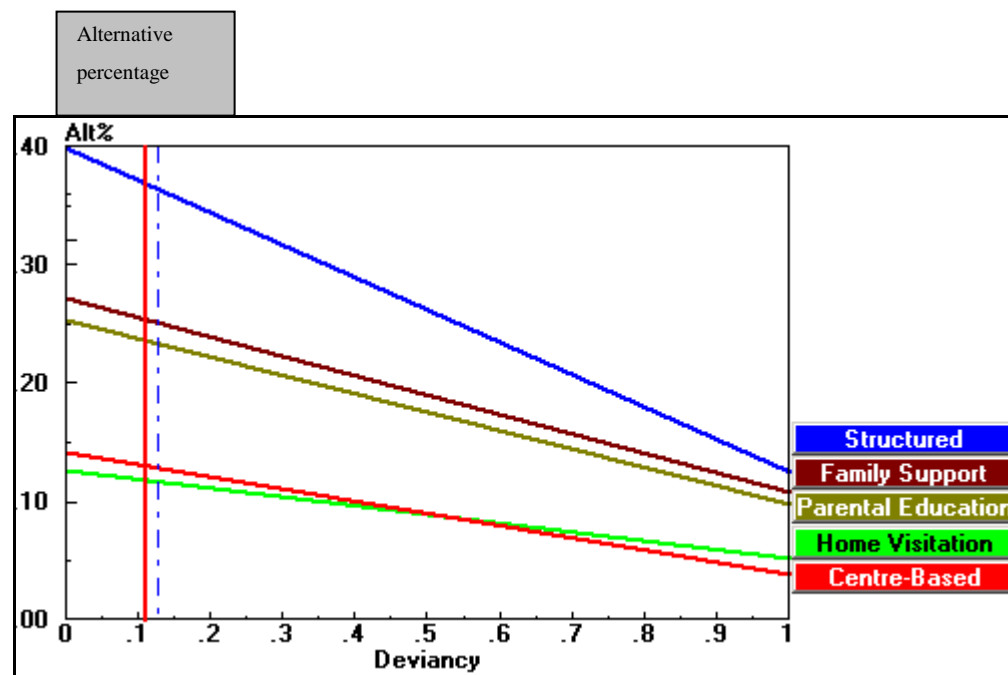
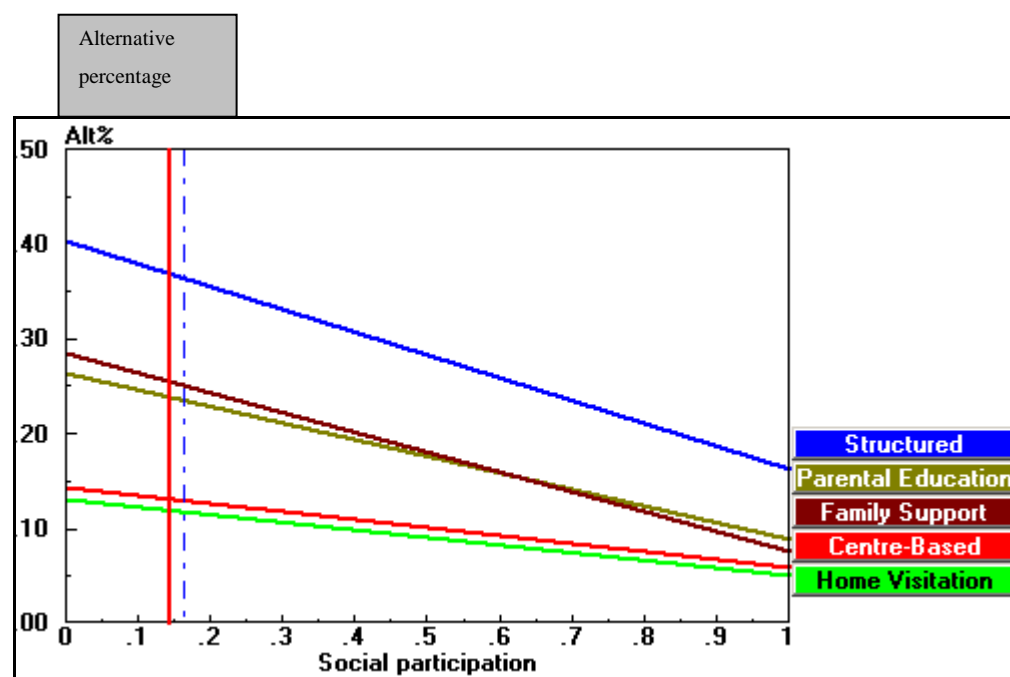


Figure 10.25 highlights that an increase in overall percentage priority outcome social participation will have the potential to change overall priority rankings of the program alternatives. The figure shows that a reduction in the overall percentage

priority rating of FW (10% in sensitivity scenario 1) has an effect on the outcome social participation, thereby increasing the overall percentage priority of social participation to 16.5 percent (an increase of 2.1%). Furthermore, Figure 10.25 shows that the outcome social participation would have to increase its overall percentage priority ranking by approximately 35 percent for any change to occur in the ranking of program alternatives family support services (FSS) and parental education (PE).

**Figure 10.25: Gradient sensitivity graph for sensitivity analysis 1, with respect to changes in the outcome SP and possible changes in program priority rankings**

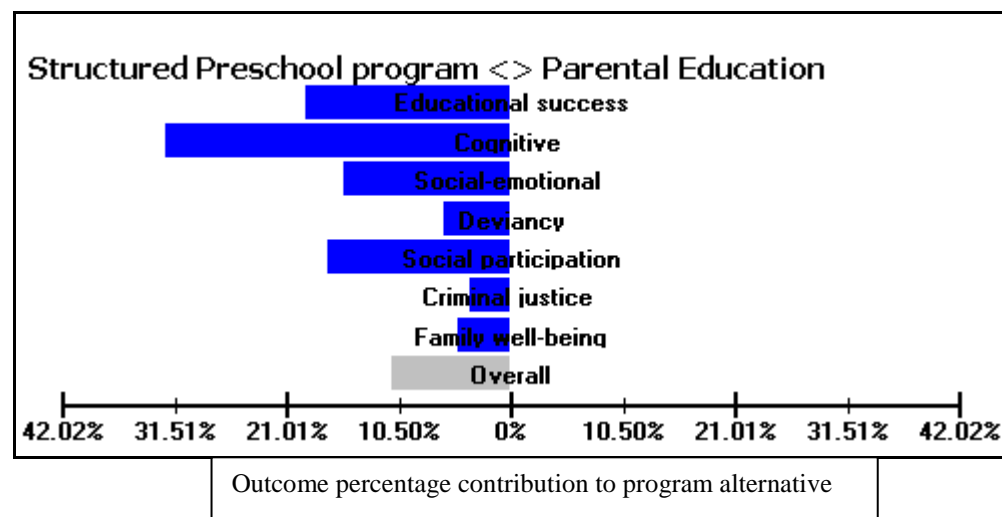


Figures 10.26 to 10.29 provide results of the first ‘what if’ scenario demonstrating how program alternatives compared to one another against the outcome domains in the decision hierarchy (in this case SPP programs vs. CBCC, HV, FSS and PE). In the figures, one early childhood intervention program is listed on the left side of the graph and the other on the right. Down the middle are the outcome objectives in the hierarchy. In these figures, if the left-hand alternative is preferred to the right with

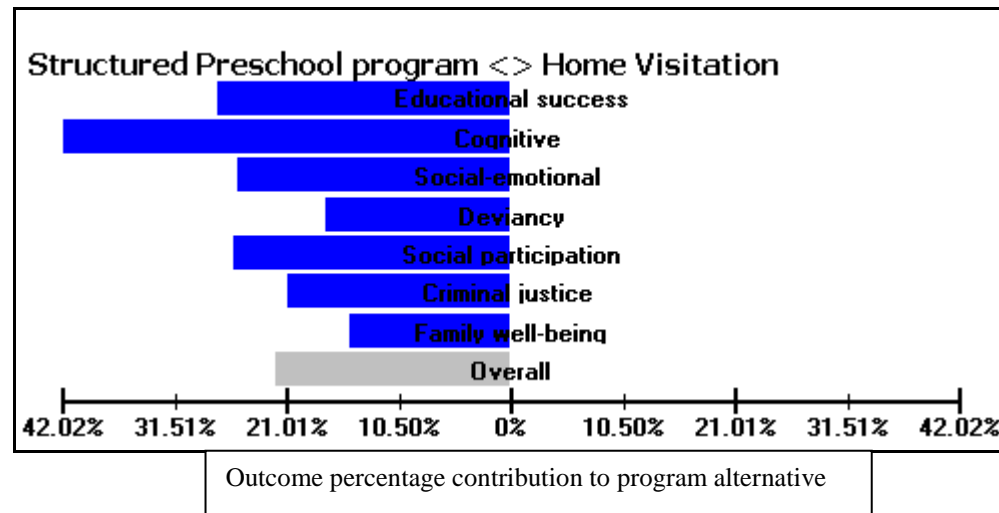


respect to an objective, a horizontal bar is displayed towards the left. Similarly, if the right hand alternative is preferred to the left, a horizontal bar is displayed to the right. If the choices are equal, no bar is displayed. At the bottom of each figure the overall result, in the form of a percentage, shows overall how much one alternative is preferred to the other. This may be seen as the composite difference between the two alternatives with respect to the seven outcome objectives.

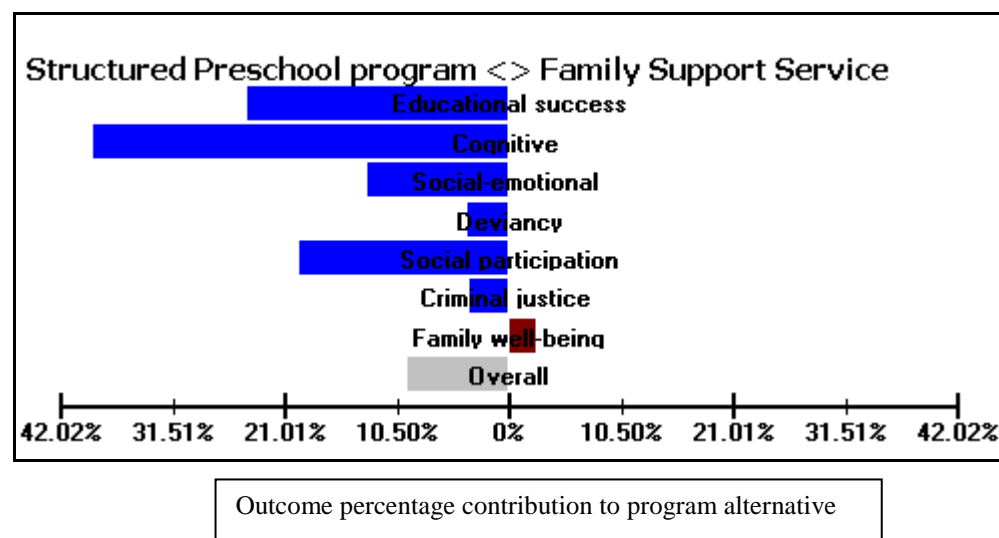
**Figure 10.26: Head to head graph of SPP vs. CBCC on seven outcome domains ('what if' scenario 1)**



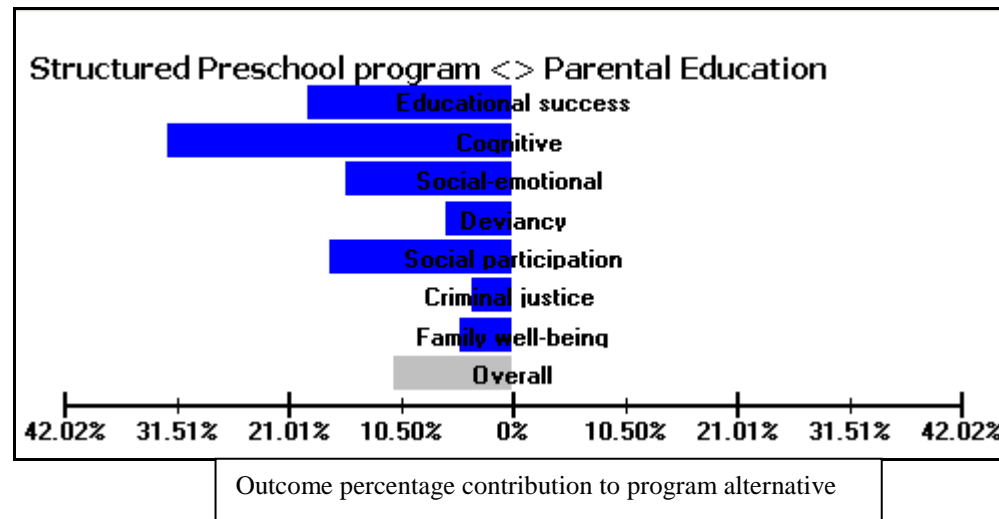
**Figure 10.27: Head to head graph of SPP vs. HV on seven outcome domains ('what if' scenario 1)**



**Figure 10.28: Head to head graph of SPP vs. FSS on seven outcome domains ('what if' scenario 1)**



**Figure 10.29: Head to head graph of SPP vs. PE on seven outcome domains ('what if' scenario 1)**



Figures 10.26, 10.27, and 10.29 highlight that SPP is preferred overall to CBCC, HV and PE respectively. Moreover, the outcome objectives (ES, CD, SED, D, SP, and CJ) are given higher priorities with respect to the SPP alternative when compared to all other program alternatives (see Figures 10.26, 10.27, and 10.29 in particular). However, the program alternative CBCC is preferred to SPP with respect to the outcome domain FW (Figure 10.28).

In the second 'what if' scenario, we reduce the most preferred outcome objective (FW) by twenty percentage points. Figure 10.30 demonstrates that if we reduce the most preferred outcome objective by twenty percent there is no change to the ranking of program alternatives. In this case, the structured preschool program (SPP) increases by an overall 1.6 percent (to 33.6% of overall ranking priorities). Centre-based childcare/developmental day-care (CBCC) programs increased by 0.02 percent (11.8%), home visitation (HV) reduced by 0.01 percent (10.6%), family

support services (FSS) decline by 1.6 percent (22.4%) and parental education (PE) programs remain stable at 21.6 percent.

**Figure 10.30: Sensitivity for outcome domains and early childhood intervention alternatives (Reducing the most preferred alternative (FW) by 20%)**

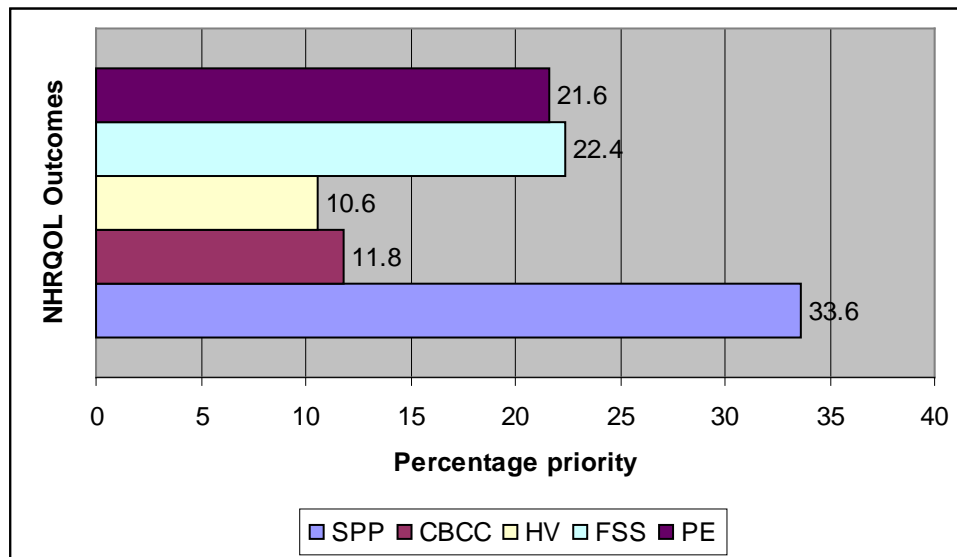
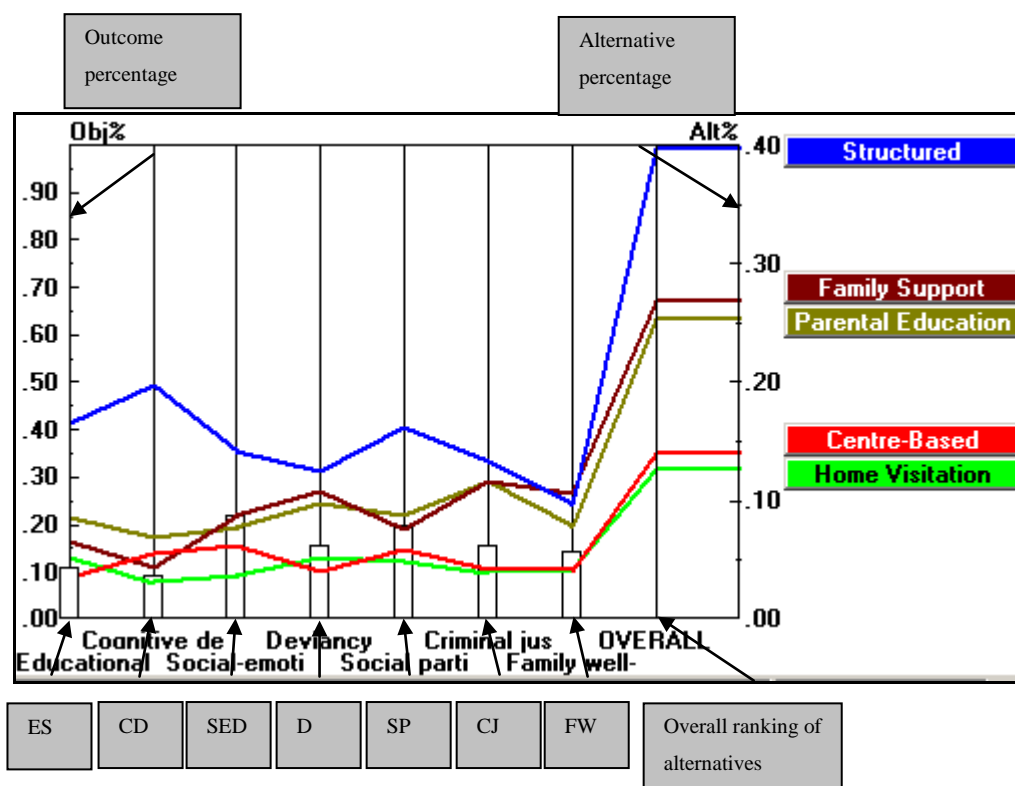


Figure 10.31 demonstrates that the early intervention program SPP rates the highest (33.6%) with respect to overall priority percentage rankings and its potential contribution to improving non health-related quality of life during adolescence. This is followed in order of percentage priority (highest to lowest) by the program alternatives FSS (22.4%), PE (21.6%), CBCC (11.8%), and HV (10.6%). Moreover, Figure 10.31 highlights that the non health-related outcome SED is rated the highest with respect to its potential contribution to improving non health-related quality of life during adolescence (20.8% of total priority); followed in order of percentage priority by the non health-related outcomes SP (18.6%), D (14.6%), CJ (14.5%), FW (13.3%), ES (10.1%), and CD (8.1%).

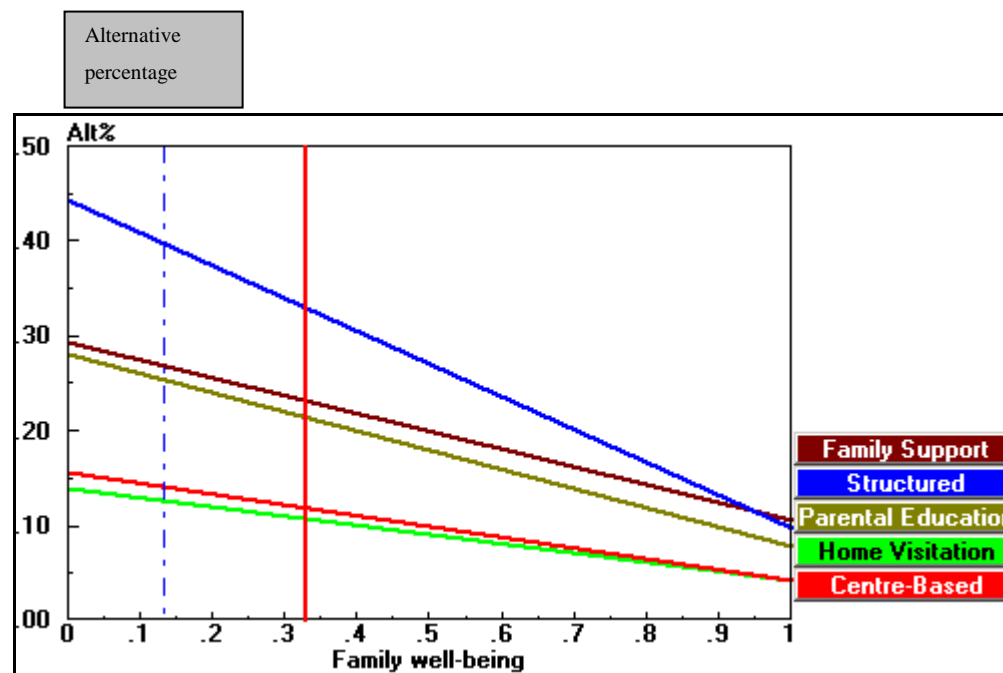
**Figure 10.31: Overall performance graph for hierarchy with a 20 percent reduction in the outcome FW**



Furthermore, Figure 10.31 allows us to identify what outcomes contributed the most to the percentage priority rankings of the five early childhood intervention program alternatives. This is represented by the coloured lines showing the overall contribution made by non health-related outcomes, in terms of their percentage of priority, to each program alternative. The coloured lines demonstrate a large gap between the outcomes ES, CD, SED and SP, whereby the blue line (representing structured preschool programs) is separated from the red, brown, khaki, and green lines. Consequently, this shows us that SPP programs received large priority percentage ratings with respect to the non health-related outcomes ES, CD, SED and SP. However, the gap was not as significant as demonstrated in Figure 10.16. Furthermore, we can see that the red (CBCC program) and green (HV program) lines separate from the other coloured lines with regards to the non health-related outcomes deviancy (D), criminal justice outcomes (CJ), and family wellbeing (FW). Thus, the outcomes D, CJ, and FW received rather low priority percentage rankings with respect to their overall contribution to the program alternatives CBCC and HV. Consequently, these outcomes influenced the overall percentage of priority rankings between CBCC and HV with respect to the program alternatives SPP, FSS, and PE.

Figure 10.32 illustrates the alternative priorities for the program alternatives (SPP, CBCC, HV, FSS and PE) with respect to the outcome objective FW, based on results of our second sensitivity analysis.

**Figure 10.32: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome FW and possible changes in program priority rankings**



Similar to Figure 10.19, we are looking for the cross-over points of the program alternatives, which indicates the direction and strength of change (in this case related to the outcome FW) required to alter the percentage priority rankings of the five early childhood intervention program alternatives. The figure demonstrates that we are a significant distance from any of these cross-over points whereby the ranking of program alternatives would change. Moreover, the outcome FW would have to have a percentage priority ranking of approximately 91.5% for a structured preschool program to be rated lower than a family support program with respect to its potential contribution to non health-related quality of life during adolescence.

Figure 10.33 demonstrates that changes in the percentage of priority ranking on the outcome domain criminal justice outcomes will not change the overall priority rankings of the five early childhood intervention programs alternatives. Further, the figure shows that a reduction on the overall percentage priority rating of CJ (20 % in sensitivity scenario 2) has an effect on the outcome CJ, thereby increasing the overall percentage priority of criminal justice outcomes to 14.5 percent (an increase of 3.3%).

**Figure 10.33: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome CJO and possible changes in program priority rankings**

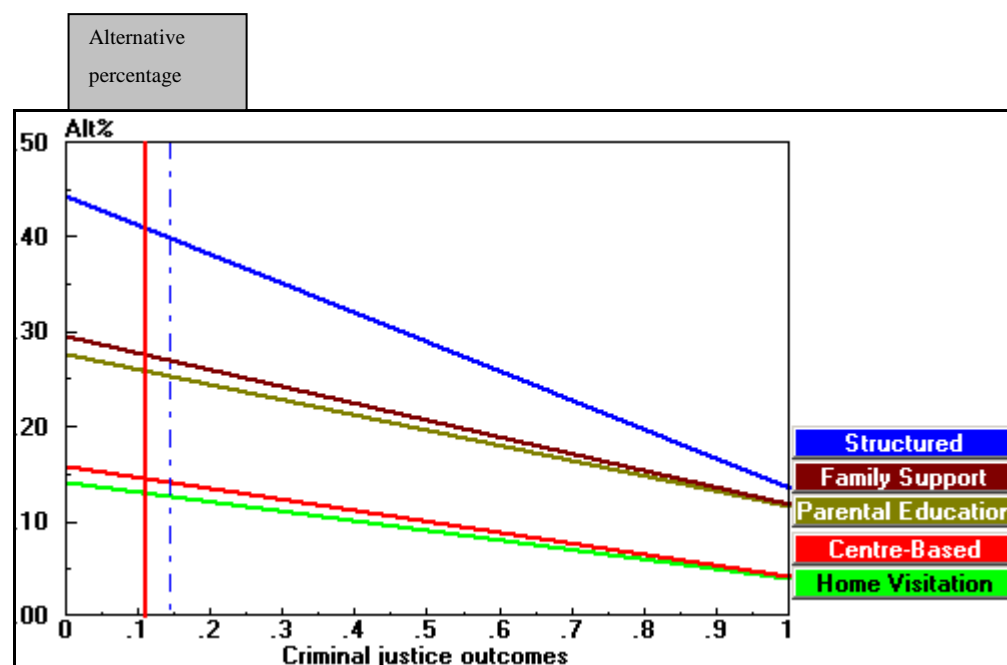




Figure 10.34 highlights that an increase in overall percentage priority outcome social participation will have the potential to change overall priority rankings on the program alternatives. The figure shows that a reduction on the overall percentage priority rating of FW (20% in sensitivity scenario 2) has an effect on the outcome social participation, thereby increasing the overall percentage priority of social participation to 18.6 percent (an increase of 4.2%). Furthermore, Figure 10.34 shows that the outcome social participation would have to increase its overall percentage priority ranking by approximately 30 percent for any change to occur in the ranking of program alternatives family support services (FSS) and parental education (PE).

**Figure 10.34: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome SP and possible changes in program priority rankings**

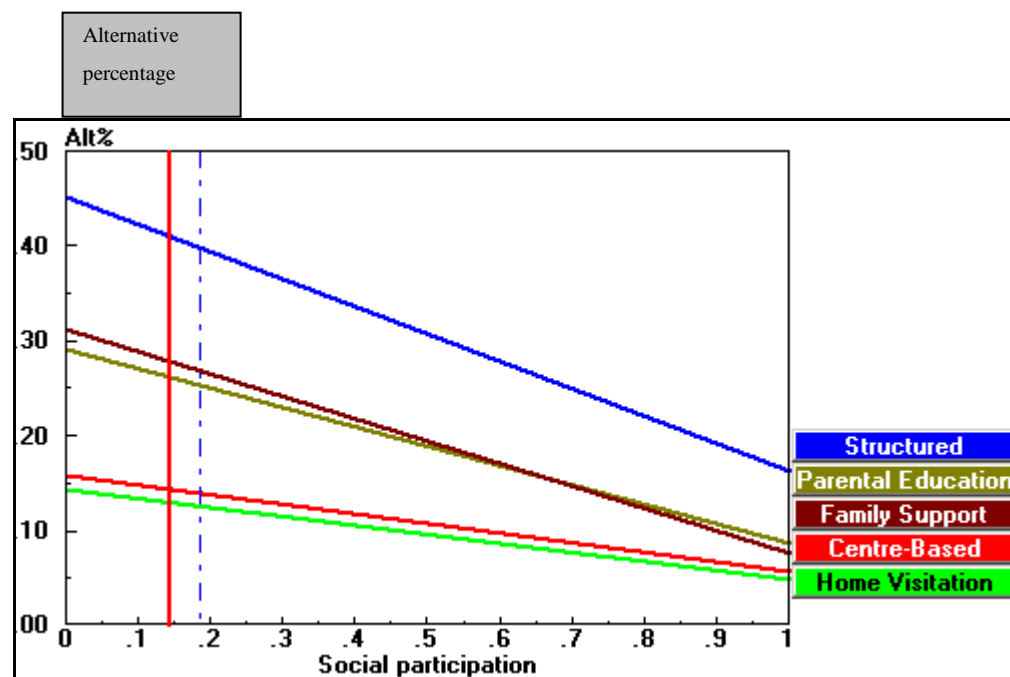


Figure 10.35 highlights that an increase in overall percentage priority outcome deviancy will have the potential to change overall priority rankings on the program alternatives. The figure shows that a reduction on the overall percentage priority rating of FW (20% in sensitivity scenario 2) has an effect on the outcome deviancy, thus increasing the overall percentage priority of deviancy to 14.6 percent (an increase of 3.3%). Furthermore, Figure 10.35 demonstrates that the outcome deviancy would have to increase its overall percentage priority ranking by approximately 40 percent for any change to occur in the ranking of program alternatives home visitation (HV) and centre-base childcare/developmental day care (CBCC).

**Figure 10.35: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome D and possible changes in program priority rankings**

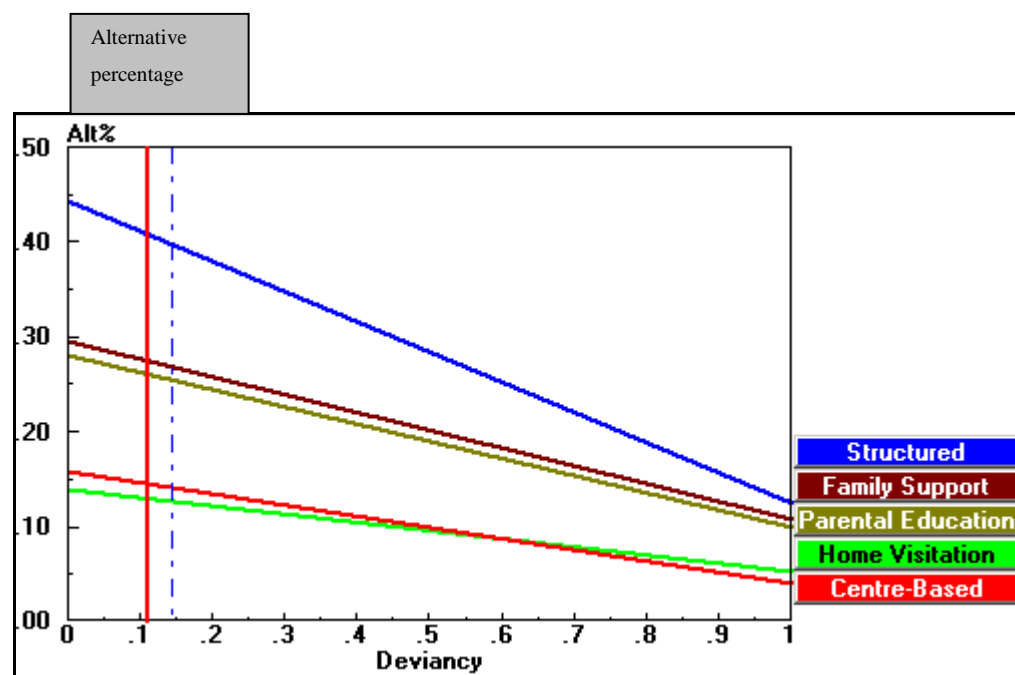


Figure 10.36 demonstrates that changes in percentage of priority ranking on the outcome domain social-emotional development (SED) will not change the overall priority rankings of the five early childhood intervention programs alternatives. Further, the figure shows that a reduction on the overall percentage priority rating of FW (20% in sensitivity scenario 2) has an effect on the outcome SED, thereby increasing the overall percentage priority of social-emotional development outcomes to 20.8 percent (an increase of 4.7%).

**Figure 10.36: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome SED and possible changes in program priority rankings**

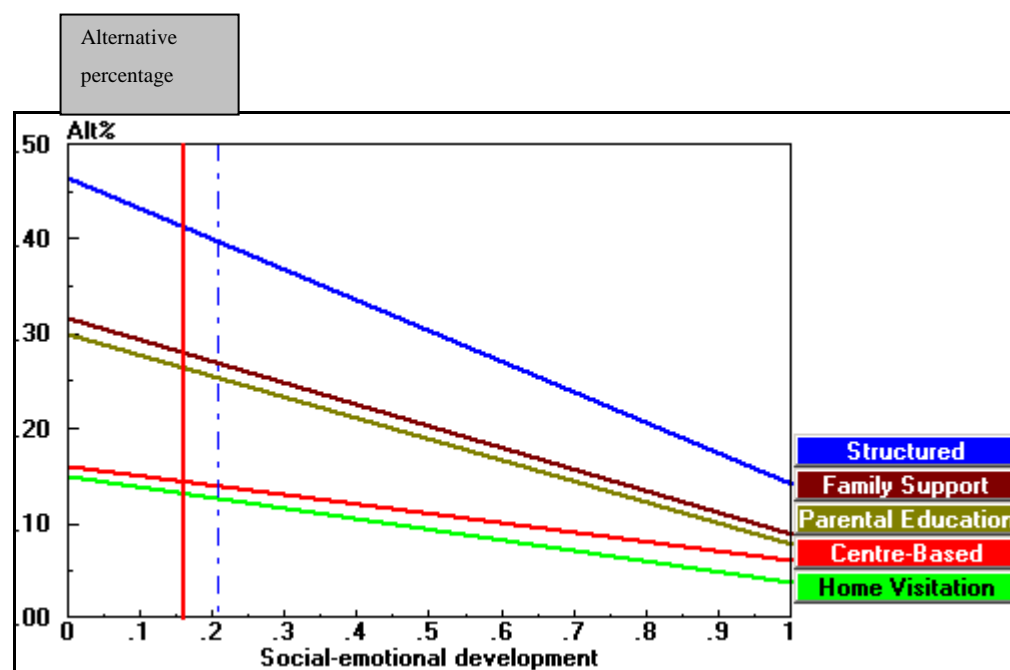


Figure 10.37 demonstrates that an increase in overall percentage priority outcome educational success has the potential to change overall priority rankings on the program alternatives. The figure shows that a reduction on the overall percentage priority rating of FW (20% in sensitivity scenario 2) has an effect on the outcome educational success, thereby increasing the overall percentage priority of educational success to 10.1 percent (an increase of 2.3%). Moreover, Figure 10.37 highlights that the outcome educational success would have to increase its overall percentage priority ranking by approximately 45 percent for any change to occur in the ranking of program alternative rankings home visitation (HV) and centre-based childcare/developmental day care (CBCC) and family support services (FSS) and parental education (PE).

**Figure 10.37: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome ES and possible changes in program priority rankings**

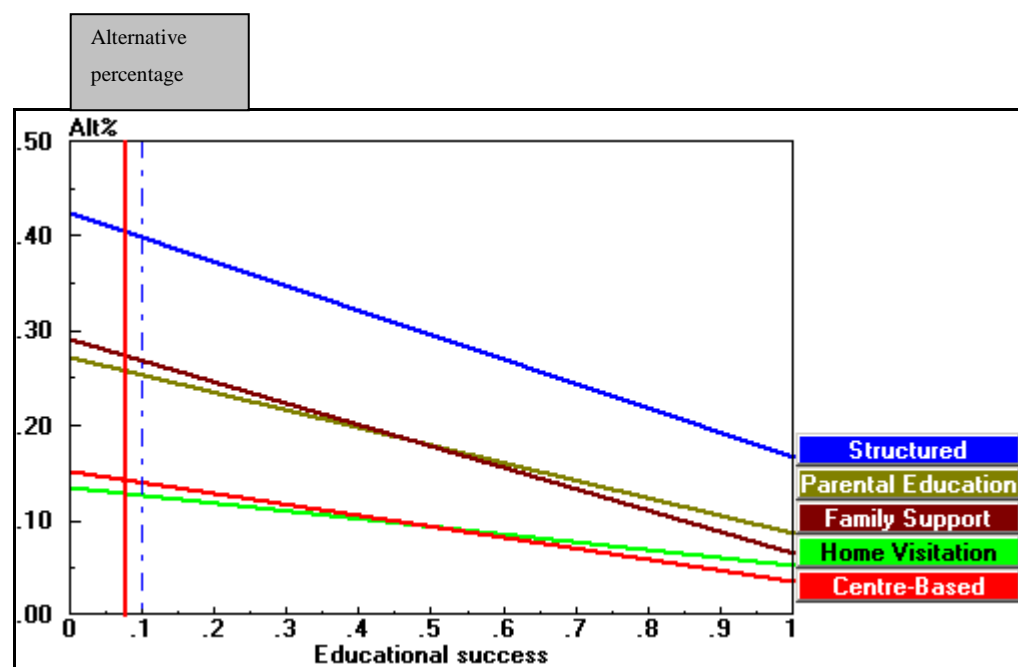
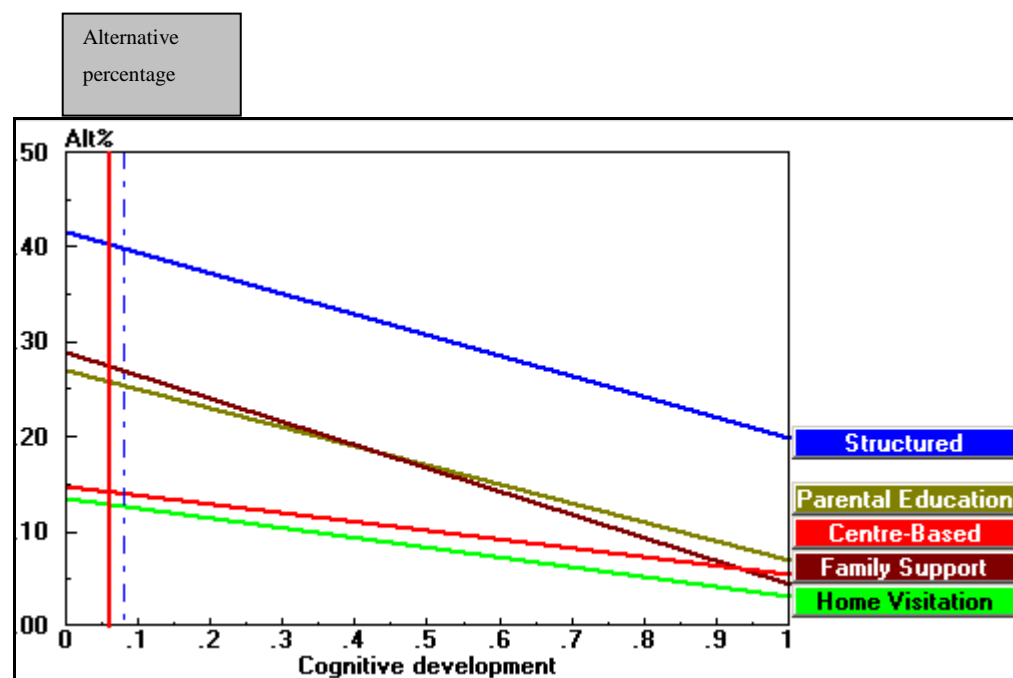


Figure 10.38 highlights that an increase in overall percentage priority outcome cognitive development will have the potential to change overall priority rankings on

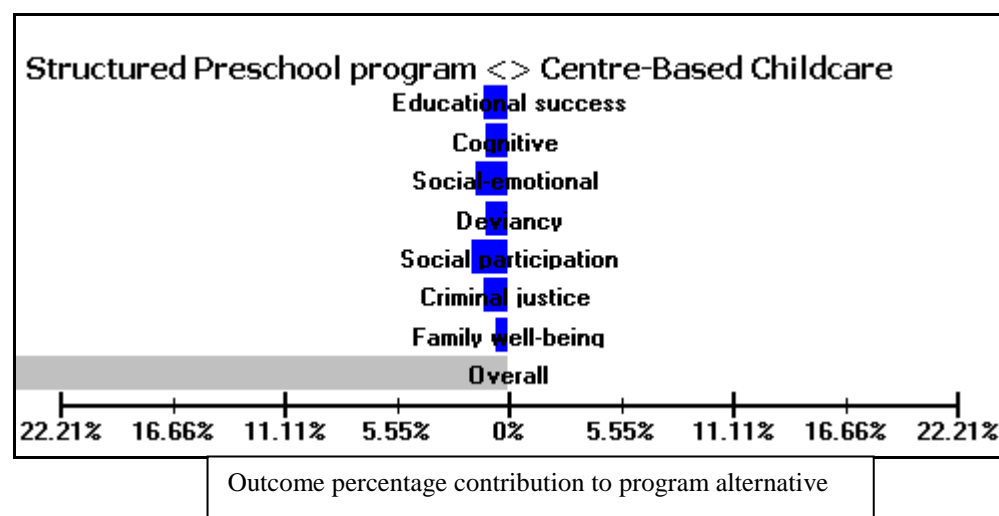
the program alternatives. The figure shows that a reduction on the overall percentage priority rating of FW (20% in sensitivity scenario 2) has an effect on the outcome cognitive development, thus increasing the overall percentage priority of educational success to 8.1 percent (an increase of 1.8%). Furthermore, Figure 10.38 demonstrates that the outcome cognitive development would have to increase its overall percentage priority ranking by approximately 35 percent for any change to occur in the ranking of program alternatives family support services (FSS) and parental education (PE). Moreover, the outcome cognitive development would have to increase its overall percentage priority ranking by approximately 90 percent for any change to occur in the ranking of program alternatives family support services (FSS) and centre-based childcare/developmental day care (CBCC).

**Figure 10.38: Gradient sensitivity graph for sensitivity analysis 2, with respect to changes in the outcome CD and possible changes in program priority rankings**

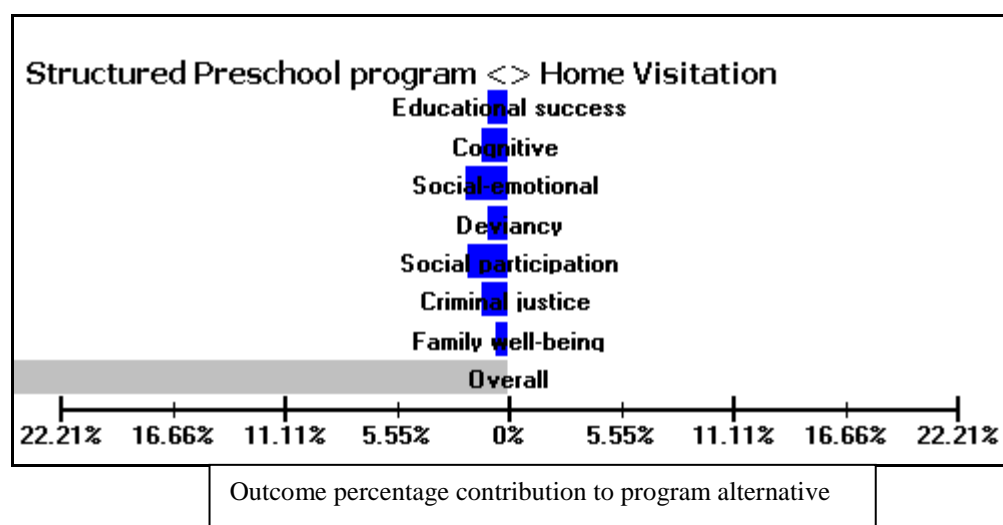


Figures 10.39, 10.40, and 10.42 highlight that under sensitivity analysis 2 assumptions SPP is preferred overall to CBCC, HV and PE. Moreover, the outcome objectives (ES, CD, SED, D, SP, and CJ) contribute more to the SPP program alternative than any other program alternatives (see Figures 10.39, 10.40, and 10.42). However, Figure 10.41 indicates that the outcome FW is considered to contribute equally to the program alternatives CBCC and SPP. Furthermore, the outcome D and CJ are considered only slightly more important with respect to the program alternative SPP when compared to FSS.

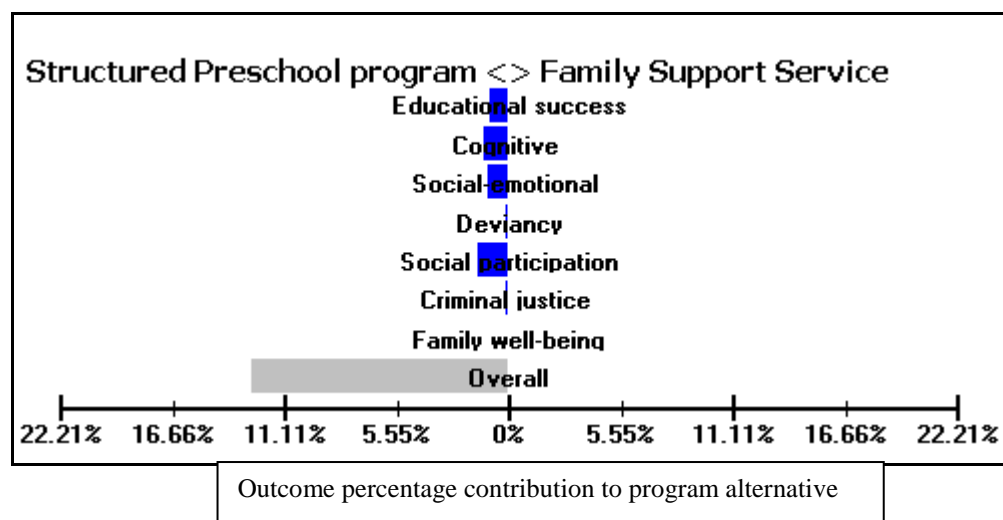
**Figure 10.39: Head to head graph of SPP vs. CBCC on seven outcome domains ('what if' scenario 2)**



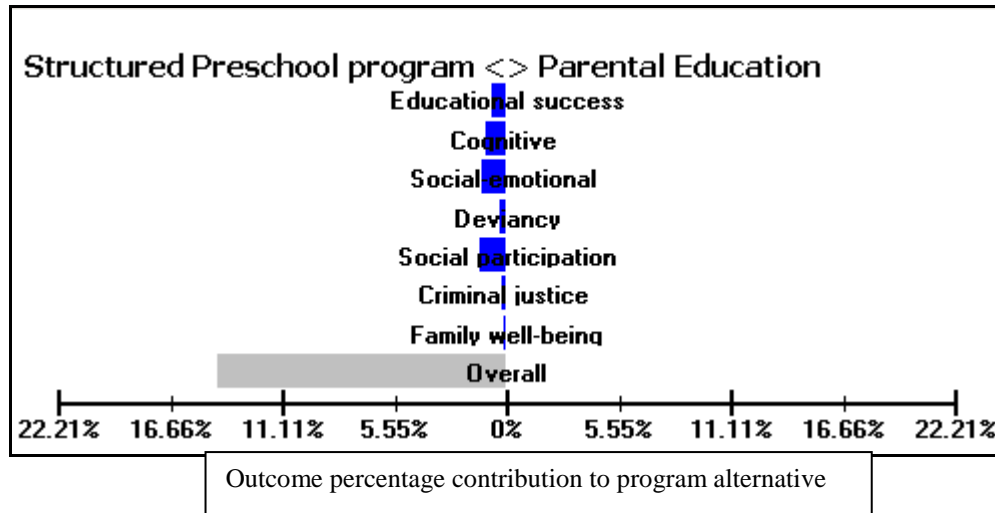
**Figure 10.40: Head to head graph of SPP vs. HV on seven outcome domains ('what if' scenario 2)**



**Figure 10.41: Head to head graph of SPP vs. FSS on seven outcome domains ('what if' scenario 2)**



**Figure 10.42: Head to head graph of SPP vs. PE on seven outcome domains ('what if' scenario 2)**



### 10.6: Discussion of results

Results of the first level of the analytic hierarchy model provided preference weights for the outcome domains ES (0.078), CD (0.06), SED (0.161), D (0.113), CJ (0.112), SP (0.144) and FW (0.330), with respect to their potential contribution to increasing non-health related quality of life (NHRQOL) during adolescence. Overall, the consistency for this matrix was 0.06, which is considered acceptable for a 7 x 7 matrix. A consistency of less than or equal to 0.10 is considered acceptable for any matrix greater than a 4 x 4. Results highlight that the outcome family wellbeing (FW) is considered the highest priority followed by social-emotional development (SED), social participation (SP), deviancy (D), criminal justice outcomes (CJ), educational success (ES), and cognitive development (CD) in descending order of priority.



This result was initially surprising given that objective research indicates that longitudinal follow-ups of early childhood interventions during adolescence only display a small effect on this outcome. However, respondents, although provided with objective research results, still considered this to be the most important outcome with respect to its potential contribution to non-health related quality of life during adolescence. This result may be, in part, due to respondents realising that some of the longitudinal early childhood intervention programs were initially developed with more emphasis placed on components such as educational success and cognitive development. Further, results of the meta-analysis were also limited given the restricted number of findings on the outcome FW during adolescence. Moreover, current views by policy-makers and the academic community may coincide with those by MacLeod and Nelson (2000), Nelson et al., (2001), Weissberg and Greenberg (1998), Yoshikawa (1994), and Zigler (1992) arguing that multi-component early intervention programs do provide positive outcomes on children's social-emotional, educational, cognitive development and improvements in family well-being. Therefore, our respondents may have been persuaded by current research (Freiberg, 2005; Homel et al., 2006) highlighting the importance of focusing not only on the at-risk child's school environment but also their home environment. Moreover, respondents may have been persuaded by research highlighting the benefits to at-risk families resulting from multi-systemic interventions. This research is important given the link between good or 'normal' family functioning and the positive effects this particular outcome has on other non health-related outcomes during adolescence (Brooks-Gunn et al., 2003).

Results from the second level of the hierarchy demonstrated that overall, respondents considered a very high effect (VH) most important when selecting an early intervention program based on its potential effect on a given outcome domain during the adolescent years. Our analysis demonstrated that the outcomes family well-being (VH-FW) (0.416) and social-emotional development (VH-SED) (0.151) are considered the highest priority with respect to their importance to contributing to NHRQOL during the adolescent years. The outcomes social participation (VH-SP) (0.115), (VH-CJ) criminal justice outcomes (0.109), (VH-D) deviancy (0.094), (VH-ES) educational success (0.067) and (VH-CD) cognitive development (0.048) constitute the remaining outcomes in order of preference.

Discussions with the participants immediately after the survey provided insights into one possible explanation for this result. Although a very high effect on each outcome was considered most important when selecting an early intervention program based on its potential effect on a given outcome domain during the adolescent years, it does not mean that a lower effect is of no importance. Rather, participants chose this as an ideal potential outcome rather than necessarily a specific criterion for inclusion for potential funding. Consequently, an ordering of preference from very high to no effect occurred. One should be cautious in interpreting this result as it does not mean that if a program does not have the potential to have a very high effect on the target group it should not be considered. Moreover, we believe that the trade-off between levels of success would not occur until one compares the cost/effectiveness of achieving high levels of effect with lower levels of effect.

Table 10.20 provides a hypothetical example, whereby the costs of achieving various levels of success are compared. In this example, we can see that the cost/effectiveness of implementing an intervention program that may potentially achieve very high effects on a set of outcomes is almost double that of funding a program with an anticipated medium effect, and almost three times that of funding a program that has the potential to achieve a small effect. Ideally, we would all strive to implement programs that have a very high effect on a given set of outcomes. However, when a policy-maker is faced with the gamble of funding a program that has a high probability of achieving a small to medium effect on a given set of outcomes with a program that has a lower probability of achieving a very high effect, we believe a trade-off will occur; particularly when costs are incorporated into the decision. This of course will depend upon the outcomes we aspire to achieve and the level of importance that the decision-maker attaches to each outcome. Moreover, policy-makers also have the ethical dilemma of choosing between providing a potentially high impact intervention on a small number of children or funding a program that could potentially have a medium effect on a larger population of children.

**Table 10.20: Hypothetical cost-effectiveness scenario: Comparing the cost/effectiveness of 3 levels of effect (very high effect, medium and small effect)**

	<b>Preschool program Alternative 1 (VH-effect)</b>	<b>Preschool program Alternative 2 (M-effect)</b>	<b>Preschool program Alternative 3 (S-effect)</b>
<b>Average cost per participant</b>	\$65,202.30	\$32,320.10	\$17,413.00
<b>Overall effectiveness</b>	70.2	60.0	55.0
<b>Cost-effectiveness ratio</b>	\$928.81	\$538.66	\$316.60

Further, Figures 10.2 to 10.11 highlight that in 60 percent of cases the outcome domain educational success (ES) rated the highest preference score with respect to a higher level of success being compared with a lower level of success. Moreover, the outcome domain family wellbeing (FW) rated the lowest on 40 percent of occasions with respect to a higher level of success being compared to a lower level of success. Additionally, the outcome domain deviancy scored the second highest preference score on sixty percent of comparisons when higher levels of success were compared with lower levels of success.

Results also demonstrated that the larger the gap between level of success (e.g. VH compared to S or N) the higher preference score it received. For example, when a very high effect was compared to no effect the average preference score was 5.57, indicating a strong importance. On the other hand, the smaller the gap between the level of success (e.g. levels of success S and N and VH and H) the smaller the preference score or level of importance. Furthermore, when large effects were being compared (e.g. VH and H, VH and M, and H and M) there was only a weak importance over one effect to the other across all outcome domains (average 2.71). Results also demonstrated that on no occasion did a respondent consider any level of success of absolute importance over another or a very strong importance over another. This suggests, in part, that respondents believed that some effect was better than no effect. Consequently, when levels of effect VH, H, M, and S were compared with no effect (N) (e.g. 5.57, 5.39, 4.76, and 3.74 respectively) levels of importance rated from a weak to strong importance. However, these results do not demonstrate a very strong or absolute importance when comparing higher levels of effect to lower levels of effect.

A final synthesis of all levels of the hierarchy demonstrated that overall structured preschool programs (SPP) were considered the highest priority with respect to their potential positive contribution to non health-related quality of life during adolescence. SPP was followed in order of preference by family support services (FSS) (0.240), parental education (PE) (0.216), centre-based childcare (0.116) and home visitation (HV) (0.107). Moreover, our research highlights that the outcome family wellbeing (FW) contributes the most to the overall priority for every early childhood intervention program alternative. This is followed in overall importance by social-emotional development (SED), social participation (SP), deviancy (D), criminal justice outcomes (CJ), educational success (ES), and cognitive development (CD) respectively. Further, our results demonstrate that SPP rated the highest on all outcomes except for family well-being where it rated second to family support services (FSS). Centre-based childcare (CBCC) and home visitation (HV) rated the lowest alternatives overall, and also were typically rated the lowest with respect to each outcome objective.

This overall outcome was consistent with results from level 1 of the hierarchy. Overall, in level 1, family-wellbeing (FW), social-emotional development (SED) and social participation (SP) were considered the highest priorities in order of preference with respect to their potential positive contribution to non-health related quality of life during adolescence. Therefore, it is no surprise that programs such as SPP, FSS and PE received such high priority rankings given that these programs aim at (at least in part) improving these outcomes with respect to children and their families. This also seems to support the proposition that there is a strong link (at least from the

perspective of the survey participants) between good or ‘normal’ family functioning and the positive effects this particular outcome has on other non health-related outcomes during adolescence (e.g. educational success, cognitive development, deviancy, and criminal justice outcomes). Moreover, the ranking of the overall program alternatives appears to support our earlier claim that survey participants consider current research to be critical to their decision-making frameworks. However, is not possible to make a link to multi-contextual or multi-component programs and the effect these may have (based on subjective priority rankings) given that we did not incorporate these into the hierarchy.

### **10.7: Limitations of the study**

It is important to highlight some limitations of this study. Preferably, we would have liked to construct a more detailed hierarchy that incorporated the most relevant indicators of our seven outcome domains (ES, CD, SED, D, CJ, SP, and FW). With this information we could have provided a more detailed set of priority weights for non health-related outcomes during adolescence. Further, we would have also liked to incorporate elements such as length of program, intensity of program, follow-up programs, multi-component programs, and multi-contextual programs into our hierarchy. However, this would have made the survey very complicated, expensive and time consuming. It is argued that a series of follow-up studies could usefully be conducted using this methodology in order to capture these various elements. The number of participants we were able to access to complete our interviews under each category (e.g. policy-maker group, academic group etc) was relatively small. However, for the purposes of this study, we considered this small sample to be a representative group of the population concerned; in particular, those

individuals making policy decisions with respect to early childhood interventions. Members of each group were selected based on their experience and knowledge in this particular area. As a result of the small sample size, the weights assigned to relative importance of the outcome objectives and the program alternatives may carry some bias. Accordingly, we conducted a thorough sensitivity analysis to measure the responsiveness of the results to changes in the relative importance of outcome objectives. We found, however, that our results were very stable with respect to significant changes to outcome objectives. Future research studies with more funding than that associated with a PhD thesis could usefully be undertaken with a larger sample size in each category. Moreover, this survey was limited to constructing priority ranking and preference values for early childhood interventions using a survey participant base limited to Brisbane, Australia. Future researchers could usefully expand the survey participant base to incorporate national (or even international) perspectives from experts within the area of early childhood interventions. Further, we were also limited by our inability to provide respondents with objective research results regarding the effectiveness of early intervention programs in Australia. Mainly due to the majority of interventions conducted and evaluated overseas, in particular, the U.S.A. Consequently, outcomes derived from our meta-analysis may not be truly representative of interventions in the Australian context. This has important implications, as respondents may have paid less attention to objective research results, given results reflect the outcomes of possibly dissimilar populations, in their decisions. However until more thorough longitudinal research into Australian early intervention programs occurs, this limitation cannot be overcome.

Some bias may present due to the analyst being the administrator of the survey. Care was taken to minimise the amount of bias introduced into the interviewing process. For example, care was taken to ensure that any opinions or extraneous comments in reaction to statements made by the various respondents did not occur. Nevertheless it needs to be acknowledged that this is a source of potential criticism of this part of the thesis. Finally, it is worth noting that the research has demonstrated that the AHP method can be applied in practice and has been shown to produce sensible results.

### **10.8: Conclusion**

In this chapter we have provided the results of our surveys using the analytic hierarchy process (AHP) method. We aimed to test the usefulness of employing the AHP method with respect to analysing complex multiple criteria problems using a systematic decision-making tool to aid in the policy planning process. More specifically, in this chapter we have examined the robustness of the AHP method for attaining preference values for measuring the economic impact of early childhood intervention programs on non health-related outcomes during adolescence. Our results appear to demonstrate that the AHP method is appropriate for both attaining preference values with respect to non health-related outcomes, and also have significant potential for use as a systematic decision-making tool to aid in the policy planning process. One of the key benefits derived from application of the AHP method in measuring the potential qualitative improvements in non health-related quality of life is that it allows the analyst to measure outcomes under uncertainty while ensuring consistency of preferences with regards to transitive social ordering of preferences. Furthermore, we argue that use of this method promises a vast



improvement on the seemingly unstructured method that is currently being used to make policy decisions in the area of early childhood intervention/crime prevention.

In the final chapter, Chapter 11, we discuss the theoretical and social policy implications of the developed methodology. Further, limitations of the current thesis are examined in more detail, together with implications for future research of our analysis.

## **CHAPTER 11: SUMMARY AND CONCLUSION**

### **11.1 Introduction**

The primary goal of prevention and early intervention is to reduce vulnerability and risk. Moreover, prevention and early intervention aim to enhance protective factors with the intention of enriching the available pathways for an individual (France & Utting, 2005). Typically, this involves the provision of access to experiences and services that compensate for adverse life circumstances, disadvantage, and vulnerability (Hayes, 2006). Over the years, however, there has emerged considerable interest in developing methods that measure the qualitative improvements that result from early childhood/crime prevention programs, with the aim of making more informed policy decisions regarding their implementation.

In Chapter 1, we identified two methodological gaps in the literature with respect to the tools available to assist decision-makers in policy decisions regarding early childhood intervention programs. The first was the lack of a methodology for making well informed choices on resource allocation and structured decision-making with respect to alternative policy options for early childhood interventions, and the second was our inability to measure the economic impact of early childhood interventions on salient life outcomes associated with non-health related quality of life (e.g. educational success, cognitive development, social-emotional development) across the life-course (e.g. early childhood and adolescence). In this thesis, we applied several methods to assist in providing solutions to these methodological gaps. Methods include: a thorough analysis of psychometric test tools used to measure the impact of early childhood interventions on non health-related outcomes across the life

course; a meta-analysis to summarise results of longitudinal studies on improvements in quality of life (non health-related) during the adolescent years; and, an adaptation of the analytical hierarchy process to elicit relative utilities and provide a structured method for making well informed choices regarding resource allocation and structured decision-making.

In this final chapter, a summary of the findings from the meta-analysis (Section 11.2) and the analytical hierarchy process (Section 11.3) are provided. Directions for future research are discussed in Section 11.4. A discussion of such future directions incorporates a detailed examination of how relative utilities (elicited from the analytical hierarchy process) may be utilised to measure the cost-utility of early childhood intervention programs. The limitations of the current thesis are identified in Section 11.5. Finally, in Section 11.6, some concluding remarks regarding the policy implications of this thesis are identified.

## **11.2: Results of the meta-analysis**

Section B (Chapters 5-6) of this thesis identified outcome domains associated with early childhood interventions and their respective indicators from early childhood through to adulthood. This was achieved by conducting a thorough examination of the psychometric test tools used by psychologists to measure improvements in quality of life (e.g. improvements in educational success, cognitive development, social-emotional development and family wellbeing) resulting from early childhood intervention programs. A detailed description of the salient outcomes and their respective indicators is provided in Appendix A. Further, we conducted a meta-analysis to establish the degree of effectiveness (in terms of effect size) of five

forms of early childhood intervention (structured preschool programs, centre-based child care/developmental day care programs, home visitation, family support services and parental education) on seven outcome domains (educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes and family wellbeing) during the adolescent life phase.

Results of the meta-analysis of longitudinal studies demonstrated that early childhood intervention projects have lasting positive effects on these seven outcome domains. These were: educational success ( $d = 0.53$ ); cognitive development ( $d = 0.33$ ); social-emotional development ( $d = 0.15$ ); deviancy ( $d = 0.39$ ); social participation ( $d = 0.37$ ); criminal justice outcomes ( $d = 0.24$ ); and, family wellbeing ( $d = 0.20$ ) during the adolescent life phase. The overall size of the effects ( $d = 0.35$ ) is in the small to medium range. Converting this to a percentile, this effect size highlights that outcome for the intervention group sample exceed by 63 percent those in the control group sample. Consequently, results indicate that the greatest effect of early childhood intervention programs during adolescence is on the outcome educational success, followed in descending order of effect by the outcomes deviancy, social participation, cognitive development, criminal justice outcomes, family wellbeing and social-emotional development. Ranking of outcomes reflect the emphasis that the five forms of early intervention programs place on outcomes. Whether the emphasis is intentional can not be determined, and requires further examination.

Our meta-analysis also revealed that early childhood intervention programs that incorporate a structured preschool or centre-based educational component

provide lasting positive effects on the outcome domains of educational success (0.48) and cognitive development (0.34) throughout the adolescent life phase.

The meta-analysis revealed that programs with a follow-through component into the early primary school years (e.g. preschool to Grade 3) had strong effects on the outcomes educational success (.57) and cognitive development (.37). Our analysis clearly reveals that the effects of follow-through are present until the late adolescent period. However, consistent with Nelson and his colleagues (Nelson et al., 2003), we were also unable to determine how much more educational intervention would be required before a plateau on cognitive and educational impacts is reached.

The meta-analysis highlighted that the length of program (e.g. > 1 year but < 3 years; and > 3 years) is an important moderator of success on the domain of educational success (> 1 year but < 3 years = .47; and > 3 years = .56). However, results were not as large on the outcome domains cognitive development (> 1 year but < 3 years = .38; and > 3 years = .32), social-emotional development (> 1 year but < 3 years = .44; and > 3 years = .12), criminal justice outcomes (> 1 year but < 3 years = .37; and > 3 years = .22), and family wellbeing (> 1 year but < 3 years = .44; and > 3 years = .14). However, when taking into account ranges and sample means, we found that programs whose duration was greater than 3 years revealed larger sample means than programs that were larger than 1 year but less than 3 years. Domains that demonstrated this effect included educational success (ES), cognitive development (CD), and family wellbeing (FW). We did not however have sufficient data to measure the effects of program length on the outcome domains social participation (SP) and deviancy (D).

The meta-analysis revealed that program intensity was an important moderator of program success. We found that a program whose intensity was greater than 500 sessions demonstrated positive results on the outcome domain educational success ( $>500$  sessions = .58; and  $\leq 500$  sessions = .38). However, using weighted means, we found results to be either in the favour of  $\leq 500$  sessions or finding no difference between either levels of intensity on other outcome domains. Further analysis of the data, taking into account ranges and sample means, found that programs whose intensity was greater than 500 sessions demonstrated larger sample means than programs that were less than or equal to 500 sessions. Outcome domains that showed this effect included educational success, cognitive development, and family wellbeing. Unfortunately, we had insufficient data to measure the effects of program length on the outcome domains social participation and deviancy. Furthermore, when combining length and intensity of programs we found both to be important moderators on the outcome domains educational success (0.56), cognitive development (0.33), and deviancy (0.49). Additionally, ranges of effect sizes indicate that programs which are longer and more intense have a greater effect on outcomes than those that are shorter and less intense.

The meta-analysis did not produce evidence supporting the notion that the greater the number of components a program had, the greater its effect on outcomes during adolescence. However, we acknowledge evidence demonstrating that when the number of program components are measured during earlier life phases (e.g. early childhood) the outcome is significantly different, with results demonstrating significantly larger effect sizes (Nelson et al., 2003). Moreover, results of previous

research indicate that multi-component programs tend to have greater effect than single component programs particularly in treating or preventing conduct disorder (Foster et al., 2007). However, during the adolescent years we found only a small difference in overall effect size between single component programs (0.436) and those comprising two components (0.441).

### **11.3: Results of the analytical hierarchy process**

Section C (Chapter 7-10) of the thesis was developed to first elicit preference values (relative utilities) for non health-related outcomes associated with early childhood intervention programs during the adolescent life phase. The second aim was to provide a methodology for analysing complex multiple criteria problems that could lead to a systematic decision making tool to aid making policy choices with respect to early childhood intervention programs and their perceived contribution to increasing non health-related quality of life during adolescence.

The analytical hierarchy process (AHP) was adapted in order to achieve the above-mentioned aims (i.e. to elicit relative utility values and provide a methodology for analysing complex multiple criteria problems). Traditionally, the analytical hierarchy process uses pair-wise comparisons to elicit relative utilities and priority rankings for comparisons between elements within the hierarchy. It does this by asking participants to provide subjective responses to pair-wise comparisons. We further developed the AHP method by incorporating objective results obtained from the meta-analysis (Section B) to assist in developing the first level of the hierarchy. In particular, we provided participants with objective research results relating to this

level of the hierarchy in order to assist them when making comparisons between alternatives.

Two surveys were conducted to achieve these aims. Participants were selected on the basis of their ability to influence decisions regarding the implementation of program options, and their demonstrated expertise and experience with respect to the evaluation of the effectiveness of outcomes resulting from early childhood interventions across the life course. Participants were selected from four distinct stakeholder groups in Queensland, Australia: (1) a policy group (e.g. representatives of Queensland Department of communities, and Department of Education, Training and the Arts); (2) a school level group (e.g. school teachers and principals); (3) a community agencies group (e.g. Mission Australia); and, (4) an academic group (e.g. academic researchers who contribute to the literature relating to developmental prevention and early education).

Survey 1 ( $n = 10$ ) consisted of two parts. Part 1 of the first survey comprised 21 questions relating to the relative preference and strength of seven outcomes with respect to their potential contribution to increasing non-health related quality of life (NHRQOL) during adolescence. Results established that the outcome family wellbeing (FW) was considered the highest priority during adolescence (0.330) followed by social-emotional development (SED) (0.161). Social participation (SP) (0.144), criminal justice outcomes (CJ) (0.112), deviancy (D) (0.113), educational success (ES) (0.078) and cognitive development (CD) (0.06) made up the remaining outcomes in order of importance.



Results of the meta-analytic review were provided to all respondents prior to the survey. Analysis of results revealed that objective research results provided to participants did not appear to determine perceived preferences for the relative importance of outcomes to non health-related quality of life (NHRQOL) during adolescence (however, this was not the main point of the meta-analysis). Moreover, what respondents consider to be a high priority during adolescence (e.g. family wellbeing and social-emotional development) with respect to contributing to NHRQOL is demonstrated by objective research to have relatively small to medium effect sizes in terms of the effect of early childhood interventions on outcomes during adolescence (e.g. family wellbeing  $d = 0.204$  and social-emotional development  $d = 0.148$ ).

These seemingly contradictory results may be due to respondents acknowledging that some of the early childhood intervention programs were initially developed with greater emphasis placed on components such as educational success and cognitive development. We posit also that had the meta-analytic findings on the outcome domain family wellbeing not been as limited (in terms on number of longitudinal findings) a much larger weighted mean effect size may have been demonstrated. As demonstrated by MacLeod and Nelson (2000), Nelson et al., (2001), Weissberg and Greenberg (1998), Yoshikawa (1994), and Zigler Taussig & Black (1992), multi-component early intervention programs do provide positive outcomes on children's social-emotional, educational, cognitive development and improvements in family well-being. It also seems likely that respondents may have been persuaded by current research (Brooks-Gunn et al., 2003; Homel et al., 2006) highlighting the importance of focusing not only on the at-risk child's school environment but also

their home environment. Moreover, respondents may have been persuaded by research highlighting the benefits to at-risk families resulting from multi-systemic interventions.

Section 2 of the first survey comprised 70 questions relating to preferences among the intensities of the outcomes. Seven matrices were developed comparing five outcome levels of success (no effect, small effect, medium effect, high effect, and very high effect) with respect to each of the seven outcomes. Results demonstrated that overall, respondents considered a very high effect (VH) most important when selecting an early intervention program based on its potential effect on a given outcome domain during the adolescent years. Moreover, the outcomes family well-being (VH-FW) (0.416) and social-emotional development (VH-SED) (0.151) were considered the highest priority with respect to their importance to contributing to NHRQOL during the adolescent years. This was followed in order of priority by the outcomes social participation (VH-SP) (0.115), (VH-CJ) criminal justice outcomes (0.109), (VH-D) deviancy (0.094), (VH-ES) educational success (0.067) and (VH-CD) cognitive development (0.048).

Although a very large effect on each outcome was considered by respondents to be the most important when selecting an early intervention program, it does not mean that a lower effect size is of no importance. A close examination of results highlights that participants chose a very large effect size as an ideal potential outcome rather than necessarily a specific criterion for inclusion or potential funding. We believe that trade-off between levels of success will probably not occur until one

compares the cost of achieving higher levels of effect with the cost of lower levels of effect.

Results of Section 2 of the first survey also demonstrated that the larger the gap between levels of success (e.g. VH compared to small (S) or no effect (N)) the higher preference score it received. For example, when a very high effect was compared to no effect the average preference score was 5.57, indicating a strong importance. Conversely, the smaller the gap between the level of success (e.g. levels of success S and N and VH and high (H)) the smaller the preference score or level of importance. Furthermore, when large effects were being compared (e.g. VH and H, VH and M, and H and M) there was only a weak importance over one effect to the other across all outcome domains (average 2.71). We also found that on no occasion did a respondent consider any level of success of absolute importance over another or of very strong importance over another; suggesting, in part, that respondents believed that some effect was better than no effect. For example, when levels of effect VH, H, M, and S were compared with no effect (N) (e.g. 5.57, 5.39, 4.76, and 3.74 respectively) levels of importance rated from of weak to strong importance.

Survey 2 ( $n = 15$ ) determined the perceived preferences for early childhood intervention programs by developing seven matrices that compared the five preschool program options (structured preschool program (SPP), centre-based developmental day care (CBCC), home visitation (HV), family support services (FSS), and parental education (PE)) in pairs, with respect to the most desired attribute outcome/level of success (VH-ES, VH-CD, VH-SED, VH-D, VH-SP, VH-CJ, and VH-FW). This survey comprised a total of 70 questions.

Results highlight that a structured preschool program (SPP) is considered the highest priority (0.320) with respect to contributing to a very high effect on all outcomes during the adolescent years. This is followed in order of priority by the programs family support services (FSS) (0.240), parental education (PE) (0.216), centre-based childcare/ developmental day care (CBCC) (0.116), and home visitation (HV) (0.107). An analysis of results demonstrates that SPP is considered three times more important than HV, and just less than three times more important than CBCC with respect to achieving a very high effect on all outcomes during the adolescent years. FSS and PE programs are considered almost equal with respect to their potential contribution to achieving a very high effect on all outcomes. In turn FSS and PE programs are seen as more than twice as important as the program alternatives CBCC and HV with regard to achieving a very high effect on all outcomes.

A final synthesis of all levels of the hierarchy showed that structured preschool programs (SPP) were considered the highest priority overall with respect to their potential positive contribution to non health-related quality of life during adolescence. SPP was followed in order of preference by family support services (FSS), parental education (PE), centre-based childcare and home visitation (HV). Additionally, our research highlights that the outcome family wellbeing (FW) contributes the most to the overall priority for every early childhood intervention program alternative. This is followed by social-emotional development (SED), social participation (SP), deviancy (D), criminal justice outcomes (CJ), educational success (ES), and cognitive development (CD). Further, results demonstrate that SPP rated the highest on all outcomes except for family well-being where it rated second to family

support services (FSS). Centre-based childcare (CBCC) and home visitation (HV) rated the lowest alternatives overall, and also were typically rated the lowest with respect to each outcome objective.

Results obtained in the meta-analysis (Section B) and the analytical hierarchy process (Section C) provide useful resources for the continual development of current research in the area of economic analyses of early-in-life intervention/crime prevention. Through the application of the results of both studies, solutions to filling the two methodological gaps (identified in Chapter 1) are possible. That is, tools to make well informed choices on resource allocation and structured decision-making with respect to alternative policy options for early childhood interventions and a method to elicit relative utility values so that the economic impact of early childhood interventions on outcomes associated with non-health related quality of life throughout an individual's life-course may be measured. The logical next step is to demonstrate how one might use the results obtained from the applied methods (meta-analysis and analytical hierarchy process) to conduct economic appraisal of early childhood intervention programs. Remembering that the aim of this thesis is to measure benefits across multiple domains, at different times, yet at the individual level, the most appropriate method to employ in economic appraisal is argued to comprise cost-utility analysis. The following section (Section 11.4) provides an insight into how future research may use the relative utility values derived from the analytic hierarchy process (Section C) in cost-utility analysis of early childhood intervention programs across the life course.

## **11.4: Directions for future research**

This thesis has provided a sound methodology for attaining preference values or relative utility values for outcomes and their respective indicators in regards to improvements in quality of life resulting from early childhood intervention programs. This method provides researchers with a tool to identify a common metric outcome/s across competing and often disparate programs with the goal of measuring the economic impact (e.g. cost-utility) of early childhood interventions on non health-related outcomes during an individual's life course. But how does one actually do this? In this section, we discuss how this may be achieved using relative utility values from the analytical hierarchy process.

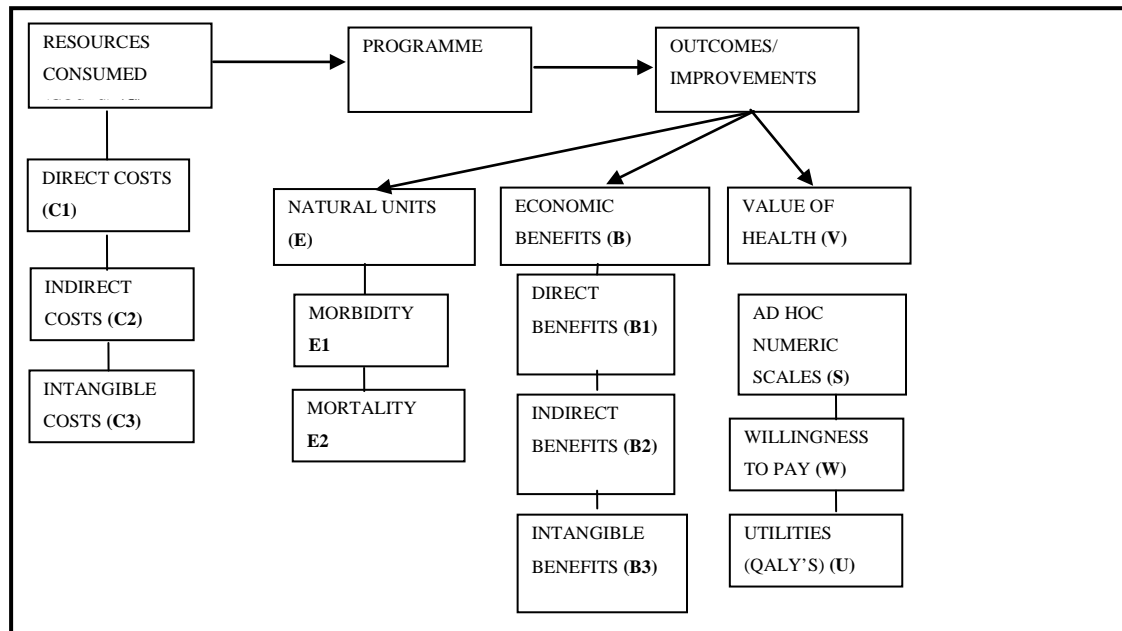
### **11.4.1: Economic appraisal of intervention programs**

The goal of any economic appraisal of a program (e.g. early childhood intervention) is to compare the amount of resources consumed (costs) with improvements created by the program (consequences) (Drummond et al., 1997). An economic appraisal consists of several components. Figure 11.1 provides an overview of the various components which include identification of both costs and consequences.

Figure 11.1 recognises that costs can be divided into three categories: direct costs (*C1*), such as personnel, facilities and equipment; indirect costs (*C2*): for example the cost of lost production because of patient participation (e.g. time off work to receive the intervention) or volunteer participation; and, intangible costs (*C3*): such as pain and suffering, grief and suffering of the participant and/or their

family, and loss of opportunities due to lack of education and the opportunities it brings (Manning, 2004).

**Figure 11.1: Components of an economic appraisal**



Source: Levin & McEwin (2001). Cost-Effectiveness Analysis, 2<sup>nd</sup> Edition, p.78

Improvements that occur as a result of a program can be measured in a number of ways (Figure 11.1). They can be measured in units that: natural to the program (*E*), such as cases prevented, or life years gained; or (b) comprise the economic benefits associated with improvements directly caused by the program (including direct benefits like savings in remedial intervention costs (*B1*); indirect benefits such as gains to society resulting from more productive individuals contributing to society (*B2*); and intangible benefits representing reduction in pain, grief and suffering of the participant and family or improved life opportunities (*B3*)); or (c) represent the value to the participant, their family and society of the possible improved outcomes (regardless of economic consequences) resulting from the intervention. A number of approaches are available to record such measures: ad hoc numeric scales (*S*),

willingness to pay or to receive (W), and perceived utility values; for example, quality-adjusted life-years, (QALY's), healthy years equivalents (HYE's), and saved young life equivalents (SAVEs), which are all based on utility measurements (Torrance, 1986). Recently, a new method has emerged in the international health policy lexicon, the disability-adjusted life year (DALY) measure (Anand & Hanson, 1996).

The QALYs approach used in health and welfare economics applies a quality-adjustment weight for different health states and then multiplies this weight by the time in a given state of health. This is then summed to arrive at a figure for quality-adjusted life years (Gold et al., 1996a). A simple example of QALYs gained from an intervention is displayed in Figure 11.2. This figure demonstrates quality of adjusted life years with and without an intervention. The shaded area between the curves represents the QALYs gained as a result of the intervention. Intuitively the process of calculating QALYs is relatively simple:

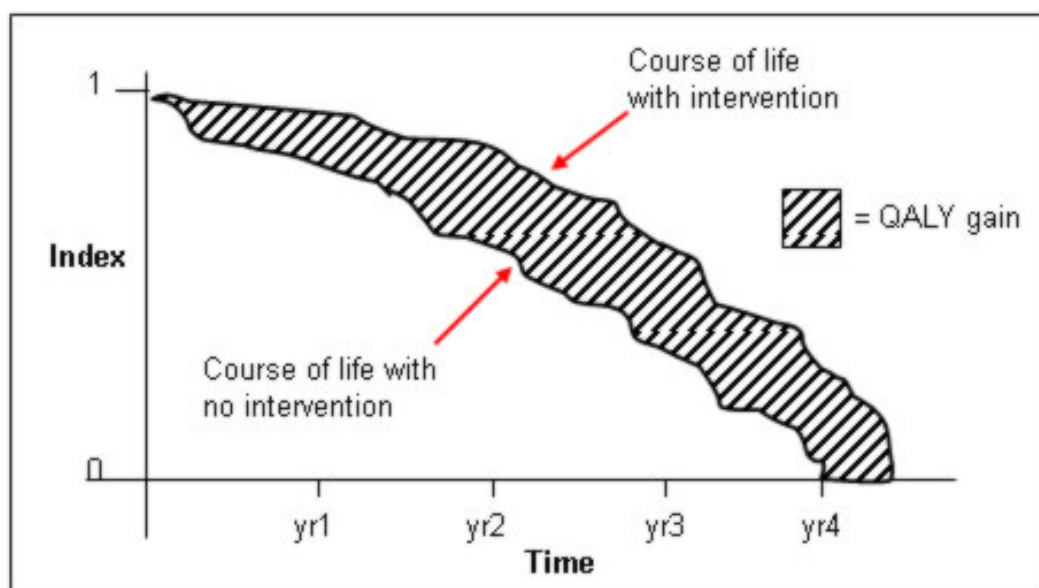
The area under the curve is simply the sum of the quality weights for the various health states on the curve (the path) multiplied by the duration (in years or fractions of years) of each health state...Because individuals and society prefer gains of all types, including health gains, to occur earlier than later, future amounts are multiplied by a discount factor to adjust for this time preference (Gold et al., 1996a, p.92-93).

Quality of adjusted life years (QALYs) is not the only method available to economists to assess the cost-utility of interventions. In fact, the QALYs methodology has been the subject of numerous criticisms. These criticisms range from those who



argue that the method is too simplistic (Gafni & Birch, 1995; Mehrez & Gafni, 1989, 1993), to those who claim that the method is over-complicated (Cox, Fitzpatrick, Fletcher, Gore, Spiegelhalter, & Jones, 1992). As a result of this, two alternatives are commonly advocated, Healthy-Years Equivalents (HYE) and Saved Young Life Equivalents (SAVEs).

**Figure 11.2: Quality adjusted life years (QALY's) gained from an intervention**



Adapted from: Gold, M. R., Patrick, D. L., Torrance, G. W., Fryback, D. G., Hadorn, D. C., Kamlet, M. S., Daniels, N., & Weinstein, M. C. (1996). Identifying and valuing outcomes. In M. R. Gold, J. E. Siegel, L. B. Russell & M. C. Weinstein (Eds.), *Cost-effectiveness in health and medicine* (pp. 82-134). New York: Oxford University Press.

Healthy-Years Equivalents (HYE), as an alternative to QALYs, assigns preferences to lifetime health paths (Gafni & Birch, 1995; Mehrez & Gafni, 1989, 1993). “HYEs are calculated by measuring the utility for each possible health path of changing health states and converting this utility through a second measurement into its HYE” (Gold et al., 1996a, p.93). There are two components to the HYE approach. The first involves the measurement of preferences over lifetime paths, and the second,

application of a two-stage standard gamble assessment. The measure of preferences over lifetime paths, is claimed by its proponents to be “conceptually highly attractive” (Gold et al., 1996a, p.93). However, critics (Buckingham, 1993; Johannesson, Pliskin, & Weinstein, 1993) have argued that applying a two-staged standard gamble is over-complicating the issue, and at the end of the day is much the same as applying a simple one-stage time-trade-off question.

The second method, Saved Young Life Equivalents (SAVEs) was developed as a result of a criticism by Harris (1987) that QALYs was an unethical method since it valued the life of a disabled person to be less than that of a person without a disability. SAVE is determined by measuring the equivalence between health gains as a result of a given intervention, and a standard measure defined as restoring full health to a young person’s life (Gold et al., 1996a).

The distinguishing feature of the SAVE approach is that it is changes in health status that are valued; both the baseline health state (e.g., survival with or without permanent disability) enter into the valuation, rather than assigning value to health states themselves...All programs would be measured in terms of their equivalent SAVE’s, and this would be the common metric of program output, replacing the QALY (Gold et al., 1996a, p.94).

The final method, the disability adjusted life year (DALY), is described by Murray, Salomon, Mathers and Lopez (2002) as the sum of the years of life lost due to premature mortality (YLL) in the population, and the years lost due to morbidity or disability (YLD) for cases of the particular health condition of interest. Simply,

DALY is a measure for the overall burden of disease. Originally developed by the World Health Organisation (Murray, 1996), DALY is now frequently used in the fields of public health and health impact assessment (Murray, 1996). DALY quantifies the impact of mortality/premature death and the disability/ impact of morbidity on a given population by combining both measures into a single comparable or common metric measure (World Health Organisation, 2007). Essentially, "...the DALY is a health gap measure that extends the concept of potential years of life lost due to premature death (PYLL) to include equivalent years of 'healthy' life lost in states of less than full health, broadly termed disability" (World Health Organisation, 2007, p.1). One DALY represents one lost year of 'healthy' life and the burden of disease as a measure of the gap between current health status and perfect, disease/disability free health status across the life course (Field & Gold, 1998). A DALY is calculated as the sum of years of life lost resulting from premature death (YLL) and the years lost due to disability (YLD); thus,  $(DALY = YLL + YLD)$ . YLL represents a given population and YLD represents incidents or cases with respect to a particular health condition (Murray et al., 2002). YLL relates to the number of deaths multiplied by the standard life expectancy at the age in which death occurs. When estimating YLD the number of incident cases in that period are multiplied by the average duration of the disease and a weight factor that relates to the severity of the disease on a scale from 0 (perfect health) to 1 (death) (Murray et al., 2002).

#### **11.4.2: Applying a cost-utility approach**

The problem with using QALY, HYE, SAVE or DALY to measure the utility derived from early childhood/crime prevention interventions is that these techniques

are specifically designed to measure impact on health related outcomes. Although it could be argued that morbidity and mortality should be incorporated into economic analyses of early childhood/crime prevention interventions, we are limited by the fact that most, if not all, evaluations of early childhood interventions conducted to date lack longitudinal data that measure impacts on rates of mortality of participants.

In Chapter 1, it was argued that cost-utility analysis (CUA) is essentially a special form of cost-effectiveness analysis (CEA). The advantage of CUA over CEA is that it measures benefits of programs using a common unit (namely utility), thus allowing programs that are dissimilar (in terms of directly comparable outcomes) to be measured and compared. As stated in Chapter 1, outcomes evaluated via CUA may be single or multiple, are general as opposed to program specific and may incorporate a notion of value (Drummond et al., 1997). In CUA, individuals are asked to express their satisfaction with single or multiple measures of effectiveness. Once costs and utilities of the various outcomes are determined, one can identify the program alternative that provides the highest utility at the lowest cost (Levin & McEwan, 2001).

Torrance (1986) argues that CUA is appropriate when you wish to employ a common unit of outcome. Further, he posits that CUA is an appropriate economic methodology when “the programmes being compared have a wide range of different kinds of outcomes, and when you wish to compare a programme to others that have already been evaluated using CUA” (p.5). Torrance, is specifically relating this statement to health-related outcomes, however, Levin and McEwan (2001) propose that the CUA approach is also an appropriate methodology to use for non health-

related outcomes. They argue that because CEA can only be used to compare the costs and typically a single measure of outcome effectiveness, an alternative method is required that can capture, or fully describe all salient outcomes that result from the intervention. Levin and McEwin use as an example the comparison of two primary school policies that (1) reduce class sizes and (2) provide out of class tutoring to improve achievement in both mathematics and reading. In this example, they state that an analysis of the two policies (from a cost-effectiveness perspective) reveals that reducing the class size is the most cost-effective means of improving reading achievement, while implementing out of class tutoring is more cost-effective in terms of improving children's mathematics scores. Given this outcome, Levin and McEwin argue that the only means of choosing between the alternatives is to estimate their relative utility.

#### **11.4.3: Measuring utility values**

In Chapter 7, we demonstrated that utility can be argued to represent a concept able to be used in describing the relative strength and preference/satisfaction that an individual (typically those directly affected or those who have an impact on an outcome) has for all plausible outcomes within a range of possibilities (Levin & McEwan, 2001). But how can we measure such relative utilities? To summarise the argument in Chapter 7, Drummond, O'Brien, Stoddart and Torrance (1997) argue that analysts have three choices in determining utility values: (1) estimating values using expert judgement; (2) looking for published values in the literature; or, (3) measuring the values using rating scales, standard gamble or time-trade-off techniques.

The first method has the advantage of being fast and inexpensive. However, as stated by Torrance (1986), “if the analysis shows that the results are sensitive to utility values, one would want to obtain utilities that are more credible, either from the literature or by measurements” (p.9). Method 2, looking for published values in the literature, has the benefit of being fast. However, it has several shortcomings. First, most published values tend to be in the health domain (e.g. Churchill, Morgan, & Torrance, 1984; Kaplan, Bush, & Berry, 1976; Torrance, Boyle, & Horwood, 1982; Torrance, Furlong, Barr, Zhang, & Wang, 1996b), with no or very few examples of non health-related utility values. Secondly, assuming that health-related utilities are what one seeks, one must ensure that the subjects used to measure those utilities are appropriate for your own study and that the instrument used to measure the utility value is credible (Torrance, 1986). Measuring values using rating scales, standard gamble or time-trade-off techniques (Method 3) is the most accurate way of obtaining values. This method however, is very costly and time consuming. Assuming one has little alternative, for example if one wishes to measure non health-related utility values, it is necessary to identify the various non health-related states for which utilities are sought, prepare descriptions of the various states, select the subjects to survey, and use one of the utility instruments described in Chapter 7 (e.g. rating scales). In the analysis in Chapters 9 and 10, preference values, or relative utilities were derived by adopting the analytic hierarchy process.

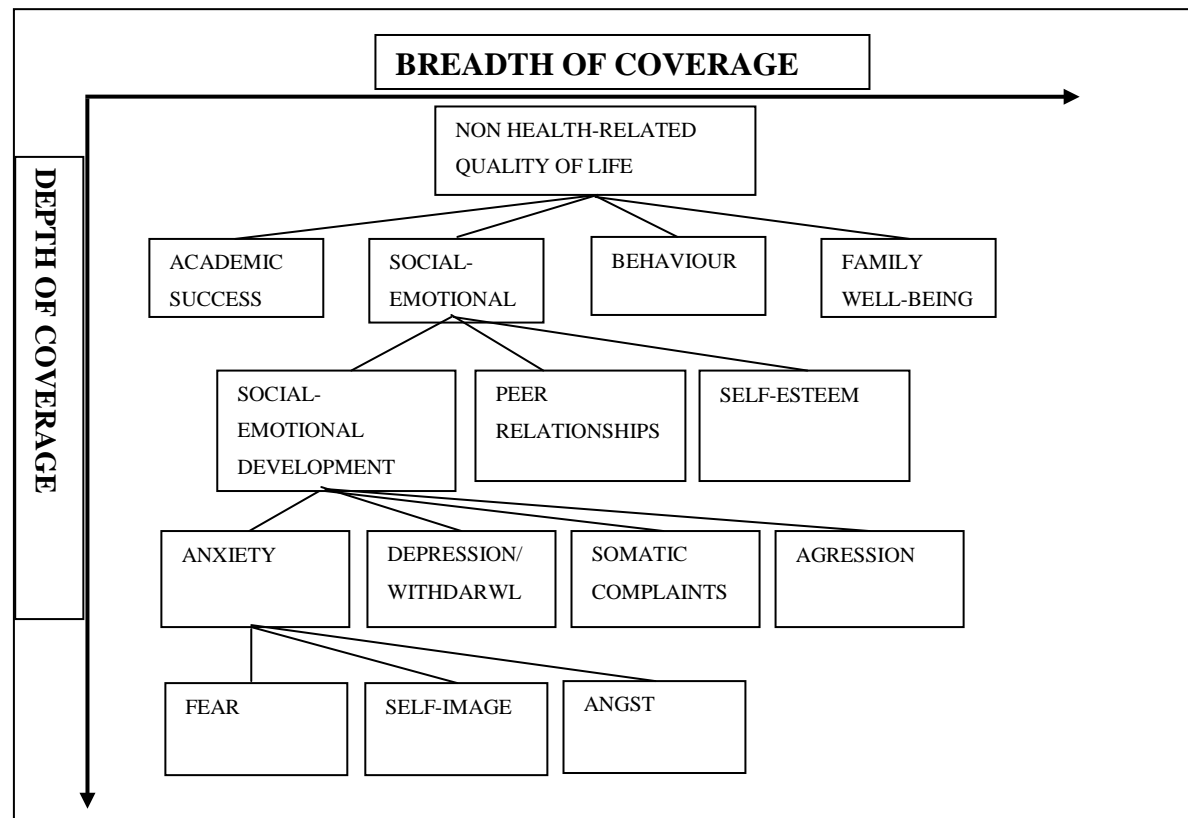
Each unique outcome, whether it is health or non health-related, for the program being evaluated and for all alternative comparison programmes must be defined as “a health state” or non health-related state “for utility measurement”

(Torrance, 1986, p.11). This may involve only a few states or hundreds of states depending on applicable outcomes. An alternative approach to describing each health or non health-related state of interest is to define a classification system that can broadly define all states of interest.

Commonly known as a multiattribute classification systems, these systems are based on the concept that health and non health-related states can be defined in terms of a number of attributes that are typically hierarchically nested (Feeny et al., 1995; Levin & McEwan, 2001; Torrance et al., 1996a). For example, if one wished to comprehensively measure the non health-related outcomes of early childhood intervention around the transition from primary school to high school one would include outcome variables such as academic/school success, social-emotional outcomes, behaviour in and out of school, and family well-being to name a few. An analysis of this type would be hierarchically based with outcome domains often having multiple sub-domains or levels (see Figure 11.3).

For example, the domain of social-emotional outcomes would consist of a number of sub-level domains (e.g. social-emotional development, relationship with peers, and self-esteem). Each domain would consist of a number of indicators. For example, social emotional development would include anxiety, depression/withdrawal, somatic complaints, psychological, attention deficiency, rule-breaking behaviour, aggression, social issues, and other problems.

**Figure 11.3: Example of attributes for a non health-related classification system**



Adapted from: Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5(1), 1-30.

Each lower level domain would be made up of a further number of indicators. For example, the lower level domain anxiety may be comprised of the indicators: *Fear* - the child has fears about going to school, is fearful of certain situations, places or animals, worries that he/she might think or do something bad; *Self-image* - the child feels the need to be perfect, complains that he/she is not loved, feels worthless or inferior; and, *Angst* - the child is constantly worried, is self-conscious or embarrassed, has feelings of guilt, speaks of suicide (Achenbach & Rescorla, 2000b). In Chapter 6 we discussed, in detail, methods for selecting specific preferences to be measured and combining them to represent a mathematical model of a subject's utility structure. In short, multiattribute utility theory provides a set of techniques for accomplishing two



tasks: “(1) quantifying the utility derived from individual attributes and (2) combining the utility from each attribute to arrive at an overall measure of utility” (Levin & McEwan, 2001, p.191).

#### **11.4.4: Using preference values to measure the economic impact of early childhood intervention programs on qualitative improvements in non-health related quality of life.**

Measuring preference values/utilities for all relevant outcomes resulting from early childhood intervention programs across the life course will be a very time consuming and expensive exercise, but one that is not impossible. In this section, we discuss how one should begin the process of conducting a cost-utility analysis of early childhood intervention programs. This involves identifying the salient attributes-or measures of effectiveness, appropriate methods measuring utility values, and a hypothetical example of a cost-utility analysis of early childhood intervention program alternatives. As argued by Nagin (2001), one should strive to measure the qualitative improvements in a child and his/her family’s quality of life as a result of early intervention. Consequently, one should aim to identify a common metric or set of common metric outcomes that accurately value the qualitative improvements in an individual’s life from a cost-utility perspective. Identifying a ‘common metric’ allows one to investigate the relationship between study features and study outcomes across a variety of programs. By coding the study features according to the objectives of the review, one may then transform the study outcomes to a common metric so that outcomes of various programs can be compared. One may then apply statistical methods to demonstrate the relationships between study features and outcomes across a variety of program options at different periods in the life course.

#### **11.4.5: Identifying outcomes relevant to early childhood intervention across the life course**

Given that no multiattribute classification system of non health-related outcomes of early childhood interventions exists, we set about an analysis of both psychometric test libraries and longitudinal research studies to identify salient outcome domains, outcomes relevant to those domains and their individual indicators. We took a life course perspective and identified relevant outcomes for specific life phases from early childhood to late adulthood (28 years +). We found that the psychometric test libraries were very helpful for identifying important non health-related outcomes and their indicators for the transition from preschool to primary school, and from primary school to high school. Longitudinal research studies made up the bulk of the data for the adolescent life transition (e.g. 13 – 18 years) with a few psychometric test tools (e.g. Achenbach behaviour checklist, HOME Inventory Administration Tests) utilised. Early adulthood (19-27 years of age) and later adulthood (27 Years+) outcomes and their relevant indicators were identified from longitudinal research studies (e.g. Perry Preschool Program (Scheinhart et al.,1993,,Parks,G ,2000 .Coalition for evidence-based policy,2005), Busselton Study (Cullen & Cullen,1996), Abecedarian Project (Campbell et al.,2002), Chicago Longitudinal Study (Reynolds et al., 2001), Elmira Prenatal/Early Infancy Project (Olds,D. et al. 1997, James ,M. 1995 )).

Data were then combined and analysed for each major life phase or transition point (e.g. transition from primary school to high school, transition from primary school to high school, the adolescent life phase, the transition from high school to post-secondary education or the workforce, early adulthood and late adulthood). With

guidance from a group of experienced psychologists from the School of Psychology and the School of Criminology and Criminal Justice, Griffith University, salient outcome domains were developed and outcomes within the individual domains and their relevant indicators categorised.

Outcomes of this analysis have been presented in Appendix A. Results presented in Appendix A represent the first comprehensive analysis and amalgamation of both psychometric test indicators and longitudinal research items for early childhood intervention program outcomes across the life course (from early childhood to late adulthood). These results highlight the salient outcome domains, sub-domains relevant to each domain and their individual indicators.

#### **11.4.6: Appropriate methods for measuring preference/utility values of non health-related outcomes**

In Chapter 7, we clarified the difference between the terms utility and preference. We argued that when measuring the perceived value (or importance weights) of non health-related outcomes, the most appropriate technique is a scaling method (rating scale). This means that preference values or relative utilities, not cardinal utilities are what needs to be measured. If we were seeking to measure cardinal utilities, we would most probably adopt a variable probability method (e.g. standard gamble technique); given that what we would then be seeking measures of utility under risk and uncertainty. The literature suggests that the variable probability technique is appropriate for measuring importance weights of health-related outcomes but not necessarily non health-related outcomes. (Levin & McEwan, 2001).

Accordingly, based on an analysis of the problem we were trying to address (measuring utility weights of non health-related outcomes of early childhood interventions), we decided to adopt a scaling method developed by Saaty (1980) for deriving our preference values or relative utilities. We adopted Saaty's method because it provided a solution to two problems faced by policy makers: (1) how to make more structured decisions based on program outcome/s at a given life phase (e.g. adolescence); and (2) how to determine the economic costs and benefits of decisions based on a cost-utility framework.

#### **11.4.7: Measuring the cost-utility of early childhood intervention programs – A hypothetical example**

The outcomes of early childhood intervention programs are difficult to express in a single measure of effectiveness (or attribute). Consequently, multi-attribute utility theory seems appropriate. Hypothetically comparing the utility and costs of three different forms of early childhood intervention (e.g. structured preschool program, home visitation and parental education classes) would involve the following steps:

- defining the attributes (outcome domains) by which success of alternatives would be judged, and identifying sub-domains and indicators that represent the attributes of interest;
- assigning importance weights to each attribute;
- generating relative utility scores;
- collecting cost data for each alternative;
- combining cost estimates and overall utility values; and,
- accounting for uncertainty

### *Step 1: Define attributes of interest*

Defining the attributes by which the success of alternatives will be judged is a critical process. This is particularly salient given that outcomes must be universal for all intervention programs being analysed and consensus obtained from all stakeholders on the appropriate attributes. This may be done in collaboration with a group of stakeholders, including policy-makers, academics, teachers, program management and staff and parents (Levin & McEwan, 2001). This collaborative effort is important given the complexity of measuring non health-related outcomes resulting from early childhood intervention across an individual's life course. As argued in Chapter 1, outcomes will differ according to the life phase being analysed. Once outcome domains are identified, one would need to identify sub-domains of interest and their individual indicators. This may be done through consultation with experts (e.g. academics, clinical psychologists) or drawing upon published material highlighting the appropriate sub-domains and indicators which should be identified for the given life phase. Part of the contribution of this thesis has been to identify and provide a structured resource from which analysts may identify relevant outcome domains, sub-domains and their respective indicators at different points throughout an individual's life (Appendix A).

### *Step 2: Assign importance weights to attributes*

Once the attributes (relevant outcome domains) of interest have been chosen, one must assign importance weights to each item in the given domain. For example, dependent variables for academic skills related to the transition from preschool to primary school may include: initiative, social relations, creative representation,

movement and music, language and literacy, and mathematics and science. Chapters 8 and 9 of this thesis provided methods of eliciting relative utility weights for outcome domains relevant to early childhood interventions in terms of their impact in the adolescent life phase. Moreover, Appendix A provides a detailed inventory of the relevant outcomes domains and their indicators from early childhood to late adulthood.

### *Step 3: Generating relative utility scores*

The Analytical Hierarchy Process (AHP) method used surveys to generate relative utility scores for five program alternatives; namely, structured preschool program, centre-based childcare/developmental day care, home visitation, family support services, and parental education. A hypothetical cost-utility example is provided in step 5, which includes the overall utility scores for the five program alternatives discussed in this thesis (see Table 11.3). These scores incorporate weights on outcome domains and intensity levels as discussed in Chapter 10.

### *Step 4: Collecting cost data for each alternative intervention*

Collection of associated program costs and the analysis of those costs is a key component of most economic evaluations of early childhood/crime prevention intervention programs. However, the process and methodology employed for collecting the required cost data are often inconsistent and incomplete. Manning, Homel and Smith (2006) argue that the following steps be adopted when conducting a thorough cost analysis of a early childhood/crime prevention program:

1. measure and value fixed costs (explicit costs);

2. measure and value implicit costs or in-kind costs;
3. distribute costs among stakeholders;
4. depreciate tangible capital assets;
5. categorise all expenditures and costs;
6. apply average and marginal costs; and,
7. discount costs for further analysis such as cost-effectiveness.

Foster, Dodge & Jones (2003) propose the use of three steps to measure the direct costs of an intervention: identifying the resources involved, measuring the use of the resources, and calculating a dollar figure. To be consistent with the economic principle of measuring all resources involved in delivering an intervention, explicit costs (costs that require an outlay of money, e.g. salaries of employees) as well as implicit costs (input costs that do not require an outlay of money, e.g. volunteer time) must be identified (Mankiw, 2001). Table 11.1 provides an overview of potential program costs. In the table, explicit costs are separated into fixed and variable cost categories. Fixed costs are costs that do not change when the number of participants increases, such as facilities. Variable costs are costs that may fluctuate depending on the number of participants (Foster et al., 2003).

**Table 11.1: Resources Used in the Delivery of an Intervention**

Variable (Explicit)	Fixed (Explicit)	Implicit
Personnel	Space	Parent time
Supplies	Utilities	Teacher time
Travel	Administration	Volunteer time
Incentives-parents	Equipment	Other space costs
Incentives-teachers	Training	
Participants' out-of-pocket costs		

Adapted from: Foster, Dodge & Jones 2003:80

Resources can be identified in several ways. The principal source in identifying program resources is budgets. Budgets clearly identify resources used as well as the cost of those resources. However, the identification of implicit costs requires that additional information such as parental reports, on-time use (e.g. teacher/volunteer time), or the estimation of costs of space, be used (Hornick, Paetsch, & Bertrand, 2000).

The valuing of resources is relatively straightforward when dealing with explicit costs, as the costs are naturally expressed in dollars within project accounts. However, difficulty lies in the careful allocation or disaggregation of costs between development, implementation, and evaluation phases of an intervention, as well as distinguishing costs associated with individual programs comprising the overall intervention (assuming that an intervention consists of several separate programs). Manning (2004) states:

For example, an intervention aimed at reducing antisocial behaviour and improving the literacy skills of grade one and two primary school children would probably consist of two separate interventions - a social skills program and a literacy enhancement program. All explicit costs of the intervention must be allocated between the two programs if one wishes to conduct further economic analysis. In the case of implicit costs, measuring resources involved is rather complex, as the initial unit measured will not be expressed in terms of dollars, and so requires additional information (p.66).

A detailed explanation of the various types of costs (e.g. explicit and implicit) is available from Manning, Homel and Smith (2006). Further, Manning et al. (2006)



employed a method of accounting for all inputs with the utilisation of a cost worksheet. The worksheet lists all ingredients and distinguishes who is paying for the costs of each alternative. A separate worksheet should be established for each program associated with the intervention. Table 11.2 provides an example of how these costs can be captured in a meaningful and methodical way.

Moreover, Manning et al (2006) outlines a detailed methodology for depreciating tangible capital assets, a method of classifying expenditures and costs of early childhood/crime prevention interventions for later evaluation, and discounting costs. Some hypothetical costs for the five program alternatives discussed in this thesis are provided in Table 11.3.

**Table 11.2: Worksheet for Estimating Costs**

Cost ingredients	Total cost	Cost to Investor	Cost to government agency	Cost to private organisations	Cost to children and parents
Personnel (includes all labour)					
Equipment (includes all durable items)					
Facilities (includes land, office space, parking space)					
Supplies (includes other consumables.					
Utilities can be included into this category)					
In-kind costs					
Intangible costs					
Total					
Ingredients					
Cost					
User Fees					
Cash					
Subsidies					
Net costs					

Adapted from: Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. Journal of Health Economics, 5(1), 1-30.

#### *Step 5: Combining cost estimates and overall utility values*

The costs (C) of each alternative are divided by its utility to yield a cost-utility ratio. Table 11.3 demonstrates that structured preschool program (SPP) provides the lowest cost-utility ratio.

**Table 11.3: Costs, utility and cost-utility ratio for three preschool program alternatives**

	Preschool program Alternative 1: (SPP)	Preschool program Alternative 2: (FSS)	Preschool program Alternative 3: (PE)	Preschool program Alternative 4: (CBCC)	Preschool program Alternative 5: (HV)
<b>Average cost per student</b>	\$30,000.00	\$28,000.00	\$25,000.00	\$20,000.00	\$15,000.00
<b>Overall utility</b>	320	240	216	116	107
<b>Cost-utility ratio</b>	\$93.75	\$116.67	\$115.74	\$172.41	\$140.18

*Step 6: Accounting for uncertainty*

The final step when conducting a cost-utility analysis is to test whether the ranking of alternatives is sensitive to key assumptions of the analysis (Levin & McEwan, 2001). The key to conducting a sensitivity analysis is to identify a key or group of key parameters. In most cases, the key parameter is the importance weight that is assigned to each attribute. Consequently, it is important to gauge whether the ranking of alternatives change when alternative sets of weights are employed (Torrance, 1986). A full description of conducting a sensitivity analysis in cost-utility studies is available from Levin and McEwin (2001).

### **11.5: Limitations**

Sections B and C of this thesis reveal several limitations with respect to both our understanding of developmental prevention and its effects into, and beyond the adolescent years, and consequent flaws in our methodology, and that of others, with regard to economic evaluations of early childhood interventions and policy decision-making processes.

With respect to Section B of the thesis, the need to understand how children from various ethno-racial backgrounds differ is acknowledged. Consequently, not only do we need to customise programs to suit these variations in ethnicity to make them more beneficial to the target groups (Homel et al., 2006; The Developmental Crime Prevention Consortium, 1999), we also need to incorporate these differences into our impact evaluations. Moreover, future research would also benefit from understanding the effects of early intervention, beyond the childhood years, on community-level outcomes. Additionally, in this thesis we were limited by the number of outcomes we could incorporate in our meta-analytic review. While this was beyond our control, it does highlight the need for evaluators to incorporate more salient outcome domains and relevant indicators into their research. Further, we were also limited by our inability to make distinctions regarding the effectiveness of early intervention programs in different geographical locations. Consequently, we had to incorporate the outcomes of longitudinal research conducted predominantly in the United States of America, which is likely not to be representative with respect to adoption of identical interventions in the Australian context. This has important implications for Part 2 of the thesis, as respondents may have paid less attention to objective research results, given results reflect the outcomes of possibly dissimilar populations, in their decisions. However until more thorough longitudinal research into Australian early intervention programs occurs, this limitation cannot be overcome.

With respect to Part C of the thesis (the analytical hierarchy process), we would have liked to have had the resources to construct a more detailed hierarchy that

incorporated the most relevant indicators associated with our seven outcome domains (ES, CD, SED, D, CJ, SP, and FW). For example, indicators of the outcome educational success include: rates of special education, school graduation, school drop-out, grade retention, completed years of education, absenteeism, reduced learning problems and the feeling of belonging at school). With this information we could have provided a more refined set of priority weights (or relative utilities) for non health-related outcomes during adolescence. Further, we would have also liked to have been able to incorporate elements such as length of program, intensity of program, follow-up programs, multi-component programs, and multi-contextual programs into our hierarchy. However, this would have made the surveys required to identify the associated pair-wise comparisons very complicated, expensive and time consuming. Consequently, we propose that a series of follow-up studies be conducted using this methodology in order to identify relative utility values incorporating these various additional elements.

We also acknowledge that the number of participants we surveyed under each category (e.g. policy-maker group, academic group etc.) was relatively small. However, for the purposes of this study, we considered those surveyed to be a representative group of the population concerned; in particular, those individuals making policy decisions with respect to early childhood interventions in Brisbane, Australia. As a result of the small sample size, the weights assigned to relative importance of the outcome objectives and the program alternatives may have caused the standard errors to be larger than otherwise. Accordingly, we conducted a thorough sensitivity analysis to measure the responsiveness of the results to changes in the relative importance of outcome objectives. We found, however, that our results were

very stable with respect to significant changes to outcome objectives. Nevertheless in future research, a selection of participants from outside of Queensland and outside of metropolitan areas would make for a more representative sample with respect to an Australian perspective. Moreover, we are not necessarily limited by the perspectives of a single country. An international perspective incorporating the views of participants from a number of different countries would represent a useful extension allowing us to improve our understanding of the impact of early childhood interventions on non health-related outcomes.

Additionally, members of each group surveyed (in this thesis) were selected based on their expert experience and knowledge in this particular area of intervention. Consequently, we did not incorporate the views of children and their families, who are important stakeholders in the decision-making process. It is argued that children are not necessarily equipped to make such decisions, although one could argue that an adolescent may be. Nevertheless it is recognised that future research should incorporate the views of parents and past program participants, and direct efforts into ascertaining the appropriateness of incorporating views of adolescents into the analysis.

In this thesis, we limited outcomes to the adolescent life phase. We chose adolescence because of the richness of follow-up data available across this life phase compared to other life phases (e.g. adulthood 28+ years). Further, previous analyses highlighting the results of early childhood interventions on non health-related outcomes during adolescence have not managed to effectively measure outcome domains beyond educational success and cognitive development. Focus was also

directed on this life phase and its major transition points given the interest by government policy-makers on outcomes associated with reductions in juvenile delinquency and crime resulting from early-in-life interventions. Finally, research was limited to the adolescent years because of the complexity of gathering preference values for all life phases. Consequently, another future research direction would be to seek to measure salient health and non health-related outcomes associated with other life phases (e.g. early childhood and adulthood).

Finally, some bias may present due to the analyst being the administrator of the survey. Care was taken to minimise the amount of bias introduced into the interviewing process from this source. For example, care was taken to ensure that any opinions or extraneous comments in reaction to statements made by the respondent did not occur. In future research, assuming more research funds are available, we propose that research staff other than the main analyst be employed to administer the survey.

### **11.6: Concluding remarks**

In Chapter 1 we stated that this thesis would be significant for four reasons. First, it provides a meta-analytic overview of the outcomes associated with early childhood intervention during the adolescent life phase. Secondly, we adapt the analytic hierarchy process method to provide a systematic way to make policy decisions with respect to choices of early childhood intervention programs on outcome domains associated with the adolescent life phase. Thirdly, the thesis provides a method for eliciting individual relative utility values with respect to choosing early childhood intervention program alternatives that potentially contribute

to improvements in an individual's non health-related quality of life. Finally, we illustrate options for measuring the cost-utility of early childhood intervention programs ensuring that one identifies a common metric outcome or set of outcomes that properly values a qualitative improvement in an individual's quality of life.

Not only are these four contributions significant from a methodological perspective, they have the potential to impact on the way policy is made. For example, with respect to the first contribution, this thesis represents the first time results from longitudinal research have been summarized on such a broad range of non health-related outcomes for at-risk children and their families during the adolescent life phase. Moreover, outcome domains have been extended to include educational success, cognitive development, social-emotional development, deviancy, social participation, criminal justice outcomes and family wellbeing and their respective indicators. Further, we have improved the methodology to ensure that outcome domains measured and their indicators are grouped according to methods employed by psychologists in psychometric tests of individuals at different life phases. Consequently, policy makers now have strong evidence (based on a summary of objective research) relating to the effect of early intervention on at-risk children and their families during the adolescence life phase.

We also argued in Chapter 1 that good decisions are difficult, if not impossible, to make without a structured process for capturing all the salient elements of the decision into the decision-making framework. Moreover, we argued that limitations to our cognitive capacity restrict our ability to make these decisions, particularly if we need to incorporate multiple elements into the decision-making



framework. This is clearly articulated in Chapters 5 and 6 and Appendix A, which summarised the array of dependent variables that potentially impact on outcomes during adolescence. Consequently, we proposed that the analytical hierarchy process be adapted to provide decision-makers with a policy development vehicle that could assist in addressing complex multiple criteria problems with the purpose of making better policy decisions. Moreover, we argue that the adapted methodology could be applied to decision-making in any area of crime prevention to make policy decisions regarding program options at any stage in the life course (e.g. early childhood, childhood, adulthood).

With regards to the third significant contribution made by this thesis we outlined the difficulty of eliciting individual relative utility values using current methods, which were implicitly non health-related (Chapter 7 and Section 11. 5.6). Capturing relative utilities is important as it allows us to use an economic approach that values improvements in quality of life resulting from early childhood intervention programs. Further, we developed a method that can be used to identify a set of common metric outcomes across competing and often disparate programs so that decisions can be made which incorporate both the stakeholder perspective and a more holistic individual perspective. Consequently, our proposed method provides a unique set of methodological tools for policy-makers to make more informed choices regarding the selection of program alternatives.

The final significant contribution of the thesis is its outline of a process for utilizing the relative utility values derived from application of the analytic hierarchy process to identify the cost-utility of early childhood intervention programs. Although a cost-utility analysis was not conducted in this thesis (since our focus was on the

development of the utility measures for inclusion in such analysis and not on the cost side), we did provide a hypothetical example of how future research may conduct such analysis.

Policy has been formulated in the field of crime prevention/early intervention for too long without a careful assessment of the benefits to the individual and his/her family. Although we have moved forward in our understanding of the need to capture individual utilities, we have fallen short; often retreating to safer ground, namely the governmental perspective. Nagin (2001) provided a wonderful analogy stating:

Just as Lewis and Clark demonstrated that overland passage to the Pacific Ocean was possible, the nascent literature on valuing developmental prevention has demonstrated the feasibility and utility of such analysis. However, just as those who succeeded Lewis and Clark found better routes to the Pacific Northwest, future economic evaluations of developmental prevention should use different analytic strategies (p.380).

This thesis has outlined and addressed the obstacles that confront the economic analyst when conducting cost-utility analyses of early childhood intervention programs. Applying and further developing the economic tools we have proposed will assist policy-makers to make better decisions regarding early intervention program alternatives. Further, the methods outlined in this thesis will contribute to better understanding with respect to the economic impact of early-in-life interventions on children and family's most at risk.

## APPENDICES

Appendices A to K are available on the C.D which is attached to the back cover of this thesis. All references relating to the appendices are incorporated into the reference list at the end of this document. Information on the attached C.D includes:

- Appendix A: Outcomes across transition points.
- Appendix B: Individual effect sizes and variances for the early childhood intervention programs included in the meta-analysis.
- Appendix C: Axiomatic foundations of the Analytic Hierarchy Process (AHP).
- Appendix D: Criticism 4 of the AHP method: The principle of rank reversal.
- Appendix E: Survey information sheet measuring the utility of early childhood intervention programs using an analytical hierarchy modeling technique.
- Appendix F: Information paper to survey participants.
- Appendix G: Ethics approval.
- Appendix H: Survey 1 questionnaire.
- Appendix I: Summary statistics survey 1.
- Appendix J: Survey 2 questionnaire.
- Appendix K: Summary statistics survey 2.

## REFERENCE LIST

- 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., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioural and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101(2), 213-232.
- Achenbach, T. M., & Rescorla, L. A. (2000a). *Child Behavior Checklist for ages 1.5-5, and caregiver-teacher report form for ages 1.5-5 (CBCL 1.5-5, C-TRF)*. Florida: Psychological Association Resources, Inc.
- Achenbach, T. M., & Rescorla, L. A. (2000b). *Child Behavior Checklist for ages 6-18*. Florida: Psychological Association Resources, Inc.
- Alexander, J. M., & Saaty, T. L. (1977a). The forward and backward process of conflict analysis. *Behavioural Science*, 22, 87-98.
- Alexander, J. M., & Saaty, T. L. (1977b). Stability analysis of the forward and backward process of conflict analysis. *Behavioural Science*, 22, 375-382.
- Allais, M. (1991). Cardinal utility history, empirical findings, and applications: An overview. *Theory and Decision*, 31(1), 99-140.
- Anand, S., & Hanson, K. (1996). Disability-adjusted life years: A critical review. *Journal of health Economics*, 16, 685-702.
- Anderson, D. (1999). The aggregate burden of crime. *Journal of Law and Economics*, 17, 611-642.
- Andrews, S. R., Blumenthal, J. B., Johnson, D. L., Kahn, A. J., Ferguson, C. J., Lasater, T. M., Malone, P. E., & Wallace, D. B. (1982). The skills of

- mothering: A study of parent child development centres. *Monographs of the Society for Research in Child Development*, 47(6), 33-42.
- Aos, S. (2003). *Costs and benefits of prevention and early intervention programs for at-risk youth: Interim report*. Washington: Washington State Institute for Public Policy.
- Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Washington: Washington State Institute for Public Policy.
- Aos, S., Phipps, P., Barnoski, R., & Lieb, R. (2001). The comparative costs and benefits of programs to reduce crime. In B. C. Welsh, D. P. Farrington & L. Sherman (Eds.), *Costs and benefits of preventing crime* (pp. 149-175). Colorado: Westview Press.
- Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, 5(2), 137-144.
- Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, 53, 941-973.
- Austin, J., Clark, J., Hardyman, P., & Henry, A. (1999). The impact of 'three strikes and you're out'. *SAGE Publications*, 1(2), 131-162.
- Baker, A. J. L., Piotrowski, C. S., & Brooks-Gunn, J. (1999). The Home Instruction Program for Preschool Youngsters (HIPPY). *The Future of Children*, 9(1), 116-133.
- Barnett, S. W. (1993). Benefit-cost analysis of preschool education: Findings from a 25-year follow-up. *American Journal Orthopsychiatry*, 63(4), 500-508.

- Barnett, S. W. (1996). *Lives in the balance: Benefit-cost analysis of the Perry Preschool program through age 27*. Michigan: High/Scope Press.
- Barnett, S. W. (2000). Economics of early childhood intervention. In S. Meisels & J. Shonkoff (Eds.), *Handbook of early childhood intervention* (pp. 589-609). New York: Cambridge University Press.
- Barzilai, J., & Golany, B. (1997). AHP rank reversal, normailization and aggregation rules. *INFOR*, 32(2), 57-63.
- Beavers, W. R., & Hampson, R. B. (1990). *Successful families: Assessment and intervention*. New York, NY: W.W. Norton.
- Belfield, C. R., Milagros, N., Barnett, S. W., & Schweinhart, L. J. (2006). The High/Scope Perry Preschool Program: Cost-benefit analysis using data from the age-40 followup. *Journal of Human Resources*, 41(1), 162-190.
- Bell, D. E., & Fuqua, P. H. (1986). Perspectives on utility theory. *Operations research*, 34(1), 179-183.
- Belli, P., Bustreo, F., & Preker, A. (2005). Investing in children's health: What are the economic benefits? *Bulletin of World Health Organisation*, 83(10), 777-784.
- Belton, V., & Gear, T. (1984). "On a short-coming of Saaty's method of analytic hierarchies. *Omega*, 11(3), 228-230.
- Berg, R. L., Hallauer, D. S., & Berk, S. N. (1976). Neglect aspects of quality of life. *Health Services Research*, Winter, 391-395.
- Bergner, M., Bobbitt, R. A., Carter, W. B., & Gilson, B. S. (1981). The sickness impact profile: Development and final revision of a health status measure. *Medical Care*, 19, 787-805.
- Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, S. W., Epstein, A. S., & Weikart, D. P. (1984). *Changes lives: The effects of the Perry Preschool Program on*

- youths through age 19*. Ypsilanti, MI: High/Scope Educational Research Foundation.
- Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (1996). *Cost-Benefit Analysis: Concepts and Practice* (2nd ed.). Upper Saddle River, New Jersey: Prentice Hall.
- Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2001). *Cost-benefit analysis: Concepts and practice* (2nd ed.). Upper Saddle River, New Jersey: Prentice Hall.
- Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2006). *Cost-benefit analysis: Concepts and practice* (3rd ed.). Upper Saddle River, New Jersey: Pearson-Prentice Hall.
- Brand, S., & Price, R. (2000). *The economic and social costs of crime*. London: Home Office.
- Brantingham, P., & Easton, S. (1998). *The costs of crime: Who pays and how much?* Vancouver: The Fraser Institute.
- Breton, J. J., Bergeron, L., Valla, J. P., Lepine, S., Houde, L., & Gaudet, N. (1995). Do children aged 9 through 11 years understand the DISC version 2.25 questions? *Journal of the American Academy of Child and Adolescent Psychiatry*, 34(7), 946-956.
- Brooks-Gunn, J., Fuligni, A. S., & Berlin, L. J. (2003). *Early child development in the 21st century : Profiles of current research initiatives*. New York; London: Teachers College Press.
- Bruer, J. T. (1999). *The myth of the first three years*. New York: The Free Press.
- Buckingham, K. (1993). A note on HYE (healthy years equivalent). *Journal of Health Economics*, 12(3), 301-309.

- Cadman, D., Goldsmith, C., Torrance, G. W., Boyle, M. H., & Furlong, M. (1986). *Development of a health status index for Ontario children. Final report to Ontario Ministry of Health for research grant DM648 (00633)*. Hamilton, Ontario: McMaster University, Centre for Health Economics and Policy Analysis.
- Caldwell, B. M., & Bradley, R. H. (2001). *HOME Inventory administration manual*. Arkansas: University of Arkansas.
- Campbell, F. A., & 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., & Ramey, C. T. (1995a). Cognitive and school outcomes for high-risk African-American students at middle adolescence. *American Educational Research Journal*, 32(4), 743-772.
- Campbell, F. A., & Ramey, C. T. (1995b). Cognitive and school outcomes for high-risk African-American students at middle adolescence: Positive effects of early intervention. *American Educational Research Journal*, 32(4), 743-772.
- Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, 6(1), 42-57.
- Canning, P. M. (1991). Profiles of the abilities of preschool aged children in an isolated northern community. *Canadian Journal of Education*, 16(1), 93-104.
- Cheng, E. W. L., & Li, H. (2001a). Analytical Hierarchy Process: An approach to determine measures for business performance. *Measuring Business Excellence*, 5(3), 30-36.



- Cheng, E. W. L., & Li, H. (2001b). Information priority-setting for better resource allocation using analytic hierarchy process (AHP). *Information Management & Computer Security*, 9(2), 61-70.
- 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](http://www.chsfs.org/Child_Family_Support_Services.html)
- Churchill, D. N., Morgan, J., & Torrance, G. W. (1984). Quality of life in end-state renal disease. *Peritoneal Dialysis Bulletin*, 4, 20-23.
- Clemen, R. T. (1996). *Making hard decisions: An introduction to decision analysis* (2nd ed.). Pacific Grove, CA: Duxbury.
- 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, B. J. (1996). Is expected utility theory normative for medical decision-making? *Medical Decision Making*, 16(1), 1-14.
- Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cohen, M. (1988). Pain suffering and jury awards: A study of the cost of crime to victims. *Law and Society Review*, 22, 537-555.
- Cohen, M., Miller, T., & Rossman, S. (1994). The costs and consequences of violent behaviour in the United States: Understanding and preventing violence. In A. Reiss & J. Roth (Eds.), *Understanding and preventing violence* (Vol. 4). Washington D.C: National Academy Press.
- Cohen, M., Rust, R., Steen, S., & Tidd, S. (2004). Willingness-to-pay for crime control programs. *Criminology*, 42(1), 89-109.

- Cohen, N., & Radford, J. (1999). *The impact of early childhood intervention on later life*. Toronto: Hincks-Dellcrest Institute.
- Commonwealth Taskforce on Child Development Health and Wellbeing. (2007). *The national agenda for early childhood*. Canberra: Australian Government: Department of Families, Community Services and Indigenous Affairs.
- Consortium for Longitudinal Studies. (1983). *As the twig is bent: Lasting effects of preschool programs*. Hillsdale, NJ: Erlbaum.
- 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., Daston, C., Lanza, E., & Schatzkin, A. (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(3), 198-207.
- Coursey, D. L., Hovis, J. L., & Schulze, W. D. (1987). The disparity between willingness to accept and willingness to pay measures of value. *The Quarterly Journal of Economics*, 102(3), 679-690.
- Cox, D. R., Fitzpatrick, R., Fletcher, A. E., Gore, S. M., Spiegelhalter, D. J., & Jones, D. R. (1992). Quality-of-life assessment: Can we keep it simple? *J R Statist Soc, A* 155(3), 353-393.
- Crowe, T. J., Noble, J. S., & Machimada, J. S. (1997). Multi-attribute analysis of the ISO 9000 registration using AHP. *International Journal of Quality & Reliability Management*, 15(2), 205-222.

- Cullen, K. J., & Cullen, A. M. (1996). Long-term follow-up of the Busselton six-year controlled trial of prevention of children's behaviour disorders. *The Journal of Pediatrics*, 129(1), 136-139.
- Cummins, R. A. (2000a). Objective and subjective quality of life: An interactive model. *Social Indicators Research*, 52, 55-72.
- Cummins, R. A. (2000b). Personal income and subjective wellbeing. *Journal of Happiness Studies*, 1, 133-158.
- Cummins, R. A., & Lau, A. L. D. (2004). The motivation to maintain subjective wellbeing. In H. Switzky (Ed.), *International review of research on mental retardation: Current perspectives on individual differences in personality and motivation in persons with mental retardation* (pp. 255-301). Amsterdam: Elsevier.
- Curran, R. (1995). Justice and the threshold of educational equality. *Philosophy of Education*, 50, 239-248.
- Currie, J. (2000). *Early childhood intervention programs: What do we know?* Toronto: Joint Center for Policy Research, IL, UCLA and NBER.
- Currim, I., & Sarin, R. (1992). Robustness of expected utility model in predicting individual choices. *Organizational Behaviour and Human Decision Processes*, 52, 544-568.
- Darling, S. M. (2003). *A comparative study of the needs and sources of support of African American and white urban rural caregivers of young children with special needs*. Paper presented at the Annual Conference of the American Council on Rural Special Education, Slat Lake City, UT.
- Daro, D. A., & Harding, K. A. (1999). Health Families America: Using research to enhance practice. *The Future of Children*, 9(1), 152-176.

- Dasgupta, P. (1995). *An inquiry into well-being and destitution*. Oxford: Oxford University Press.
- Diener, E., Suh, E., & Oishi, S. (1997). Recent findings on subjective well-being. *Indian Journal of Clinical Psychology*, 24(1), 25-41.
- Dolan, P., & Gudex, C. (1995). Time preference, duration and health state valuations. *Health Economics*, 4, 289-299.
- Dolan, P., Gudex, C., Kind, P., & Williams, A. (1996). The time trade-off method: results from a general population study. *Journal of Health Economics*, 15, 209-231.
- Donohue, J., & Siegelman, P. (1998). Allocating resources among prisons and social programs in the battle against crime. *Journal of Legal Studies*, 27, 1-43.
- Drummond, M. F., O'Brien, B., Stoddart, G. L., & Torrance, G. W. (1997). *Methods for economic evaluation of health care programmes* (2nd ed.). New York: Oxford University Press.
- Drummond, M. F., Stoddart, G. L., & Torrance, G. W. (1987). *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press.
- Dunst, C. J., Jenkins, V., & Trivette, C. M. (1984). The family support scale: Reliability and validity. *Journal of Individual, Family and Community Wellness*, 1, 45-52.
- Durlak, J. (1998). Common risk and protective factors in successful prevention programs. *American Journal of Orthopsychiatry*, 68(4), 512-520.
- Durlak, J. A., & Lipsey, M. W. (1991). A practitioner's guide to meta-analysis. *American Journal of Community Psychology*, 19, 291-332.

- Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: A meta-analytic review. *American Journal of Community Psychology*, 25(2), 115-152.
- Durlak, J. A., & Wells, A. M. (1998). Evaluation of indicated preventive intervention (secondary prevention) mental health programs for children and adolescents. *American Journal of Community Psychology*, 26(5), 775-802.
- Dyer, J. S. (1990). Remarks on the analytic hierarchy process. *Management Science*, 36(3), 249-258.
- Dyer, J. S., & Wendell, R. E. (1985). *A critique of the analytic hierarchy process: Working paper 84/85*. Austin, Texas: Department of Management, University of Texas.
- Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence - The impact of Stage-Environment Fit on Young Adolescents': Experiences inschools and in families. *American Psychologist*, 48(2), 90-101.
- Eckenrode, J., Ganzel, B., Henderson, C. R., Smith, E., Olds, D., Powers, J., Cole, R., Kitzman, H., & Sidora, K. (2000a). Preventing child abuse and neglect with a program of nurse home visitation. *Journal of the American Medical Association*, 284(11), 1385-1391.
- Eckenrode, J., Ganzel, B., Henderson, C. R., Smith, E., Olds, D., Powers, J., Cole, R., Kitzman, H., & Sidora, K. (2000b). Preventing child abuse and neglect with a program of nurse home visitation. *JAMA*, 284(11), 1385-1391.
- Eckenrode, J., Olds, D., Henderson, C. R., Kitzman, H., Luckey, D., Pettitt, L. M., Sidora, K., Morris, P., Powers, J., & Cole, R. (1998). Long-term effects of

- nurse home visitation on children's criminal and anti-social behaviour. *Journal of the American Medical Association*, 280(15), 1302.
- Edelbrock, C., Costello, A. J., Dulcan, M. K., Kalas, R., & Conover, N. C. (1985). Age differences in the reliability of the psychiatric interview of the child. *Child Development*, 56(1), 265-275.
- Elder, G., Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the life course* (pp. 3-19). New York: Kluwer Academic/Plenum Publishers.
- Elder, G. H. (1998). The life course as developmental theory. *Child Development*, 69(1), 1-12.
- Expert Choice. (2000-2004). Expertchoice. Arlington, VA, USA: Expert Choice Inc.
- Eyeberg, S., & Pincus, D. (1999). *Eyeberg Child Behavior Inventory & Sutter-Eyeberg Student Behavior Inventory-Revised*. Odessa, FL: Psychological Assessment Resources.
- Eysenck, H. J. (1978). An exercise in mega-silliness. *American Psychologist*, 33, 517.
- Farrell, G., Bowers, K., & Johnson, S. D. (2004). Cost-benefit analysis for crime science: Making cost-benefit analysis useful through a portfolio of outcomes. In M. Smith & N. Tilley (Eds.), *Crime science: New approaches to preventing and detecting crime*. London.
- Farrington, D. P. (1996). Understanding and preventing youth crime. *Findings-Social Policy research*, 93, 1-4.
- Farrington, D. P., & Welsh, B. C. (2002). Family-based crime prevention. In L. W. Sherman, D. P. Farrington, B. C. Welsh & D. Layton MacKenzie (Eds.), *Evidence-based crime prevention* (pp. 22-55). London: Routledge.

- Farrington, D. P., & Welsh, B. C. (2003). Family-based prevention of offending: A meta-analysis. *The Australian and New Zealand Journal of Criminology*, 36(2), 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(2), 47-60.
- Feeny, D. H., Boyle, M. H., & Torrance, G. W. (1995). Multi-attribute health status classification systems: Health Utilities Index. *Pharmaco Economics*, 7, 490-502.
- Feeny, D. H., Torrance, G. W., & Labelle, R. (1996). Integrating economic evaluations and quality of life assessments. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 85-95). Philadelphia, New York: Lippincott-Raven.
- Field, M. J., & Gold, G. M. (1998). *Summarizing population health: Directions for the development and application of population metrics*. Washington D.C: National Academy Press, Institute of Medicine.
- Foster, E. M., Olchowski, A. E., & Webster-Stratton, C. (2007). Is stacking intervention components cost-effective? An analysis of the Incredible Years Program. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46(11), 1414-1424.
- Foster, M. E., Dodge, K. A., & Jones, D. (2003). Issues in the Economic Evaluation of Prevention Programs. *Applied Developmental Science*, 7(2), 76-86.
- France, A., & Homel, R. (2006). Societal access routes and developmental pathways: Putting social structure and young people's voice into the analysis of Pathways

- into and out of crime. *The Australian and New Zealand Journal of Criminology*, 39(3), 295-309.
- France, A., & Utting, D. (2005). The paradigm of 'risk and protection focused prevention' and its impact on services for children and families. *Children and Society*, 19(2), 77-90.
- Freiberg, K., Homel, R., & Lamb, C. (2007). The pervasive impact of poverty on children: Tackling family adversity and promoting child development through the pathways to prevention project. In A. France & R. Homel (Eds.), *Pathways and crime prevention: Theory, policy and practice* (pp. pp.226-246). Cullumpton, Devon, U.K: Willan Publishing.
- Freiberg, K., Homel, R., Batchelor, S., Carr, A., Lamb, C., Hay, I., Elias, G., & Teague, R. (2005). Pathways to participation: A community-based developmental prevention project in Australia. *Children and Society*, 19, 144-157.
- Gafni, A., & Birch, S. (1995). Preferences for outcomes in economic evaluation: An economic approach to addressing economic problems. *Social Science and Medicine*, 40(6), 767-776.
- Garber, A. M., Weinstein, M. C., Torrance, G. W., & Kamlet, M. S. (1996). Theoretical foundations of cost-effectiveness analysis. In M. R. Gold, J. E. Siegel, L. B. Russell & M. C. Weinstein (Eds.), *Cost-effectiveness in health and medicine* (pp. 25-53). New York: Oxford University Press.
- Glass, G., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Newbury Park: Sage.
- Gold, M. R., Patrick, D. L., Torrance, G. W., Fryback, D. G., Hadorn, D. C., Kamlet, M. S., Daniels, N., & Weinstein, M. C. (1996a). Identifying and valuing



- outcomes. In M. R. Gold, J. E. Siegel, L. B. Russell & M. C. Weinstein (Eds.), *Cost-effectiveness in health and medicine* (pp. 82-134). New York: Oxford University Press.
- Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (1996b). *Cost-effectiveness in health medicine*. New York: Oxford University Press.
- Golly, A. M., Stiller, B., & Walker, H. M. (1998). First Step to Success: Replication and social validation of an early intervention program. *Journal of Emotional and Behavioural Disorders*, 6(4), 243-250.
- Gordon, R. S. J. (1983). An operational classification of disease prevention. *Public Health Reports*, 98(2), 107-109.
- Grabowski, H. G., & Hansen, R. W. (1996). Economic scales and tests. In B. Spilker (Ed.), *Quality of life and Pharmacoeconomics in clinical trials* (pp. 79-95). Philadelphia, New York: Lippincott-Raven.
- Gray, S. W., & Klaus, R. A. (1970). The Early Training Project: A seventh-year report. *Child Development*, 41(4), 909-924.
- Greenwood, P., Karoly, L., Everingham, S., Hoube, J., Kilburn, M., Rydell, C., Sanders, M., & Chiesa, J. (2001). Estimating the costs and benefits of early childhood interventions. In B. C. Welsh, D. P. Farrington & L. Sherman (Eds.), *Costs and benefits of preventing crime* (pp. 123-148). Colorado: Westview Press.
- Greenwood, P., Model, K., Rydell, P., & Chiesa, J. (1996). Diverting children from a life of crime. *RAND Publication*, 1-69.
- Halpern, R., Baker, A. J. L., & Piotrkowski, C. S. (1993). *The Child Classroom Adaptation Inventory*. New York: National Council of Jewish Women.

- Hanson, R. K., & Morton-Bourgon, K. (2004). *Predictors of sexual recidivism: An updated meta-analysis*. Ottawa, Canada: Department of the Solicitor General.
- Harker, P. T., & Vargas, L. G. (1987). The theory of ratio estimation: Saaty's analytic hierarchy process. *Management Science*, 33(11), 1383-1404.
- Harker, P. T., & Vargas, L. G. (1990). Reply to "Remarks on the analytic hierarchy process" by J.S. Dyer. *Management Science*, 36(3), 269-273.
- Harris, J. (1987). QALYfying the value of life. *Journal of Medical Ethics*, 13, 117-123.
- Hawkins, D. J., Catalano, R., Kosterman, R., Abbot, R., & Hill, K. (1999). Preventing adolescent health-risk behaviours by strengthening protection during childhood. *Archives of Pediatrics and Adolescent Medicine*, 153, 226-234.
- Hayes, A. (2006). Maintaining the gains: Sustainability in prevention and early intervention. *Family Matters*, 75, 66-69.
- Heckman, J.J., Stixrud, J., Urzua, S. (2006). The effects of cognitive and non-cognitive abilities on labour market outcomes and social behaviour. *Journal of Labour Economics* 24(3), 411-482.
- Hedges, L., & 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(2), 107-128.
- Hertzman, C. (1999). Population health and human development. In D. P. Keating & C. Hertzman (Eds.), *Developmental health and the wealth of nations: Social, biological, and educational dynamics* (pp. 21- 40). New York: Guilford Press.
- Hertzman, C., & Wiens, M. (1996). Child development and long-term outcomes. *Social Science and Medicine*, 43(7), 1083-1095.

- High/Scope Educational Research Foundation. (1992). *Manual of the High/Scope Child Observation Record for ages 2.5-6*. Ypsilanti, Michigan: High/Scope Press.
- Hirshleifer, J., & Riley, J. G. (1992). *The analytics of uncertainty and information*. Cambridge, England: Cambridge University Press.
- Hockfield, S., & Lombroso, P. (1998). Development of the cerebral cortex: IX. Cortical development and experience: I. *Journal of American Academic Child Adolescence Psychiatry*, 37(9), 992-993.
- Homel, R. (2005). Developmental crime prevention. In N. Tilley (Ed.), *Handbook of crime prevention and community safety* (pp. 71-106). Cullompton, Devon, U.K: Willan Publishing.
- Homel, R., Elias, G., & Hay, I. (2001). Developmental prevention in a disadvantaged community. In R. Eckersley, J. Dixon & B. Douglas (Eds.), *The social origins of health and well-being* (pp. 269-278). Cambridge: Cambridge University Press.
- Homel, R., Freiberg, K., Lamb, C., Leech, M., Carr, A., Hampshire, A., Hay, I., Elias, G., Manning, M., Teague, R., & Batchelor, S. (2006). *The Pathways to Prevention Project: The first five years 1999-2004*. Brisbane: Griffith University and Mission Australia.
- Hornick, J., Paetsch, J., & Bertrand, L. (2000). *A Manual on Conducting Economic Analysis of Crime Prevention Programs*: Canadian Institute for law and The Family.
- Horton, N. J., Laird, N. M., & Zahner, G. E. P. (1999). Use of multiple informant data as a predictor in psychiatric epidemiology. *International Journal of Methods in Psychiatric Research*, 8(1), 6-18.

- Isard, W., & Smith, C. (1982). *Conflict analysis and practical conflict management: An introduction to peace science*. Cambridge, Massachusetts: Ballinger Publishing Company.
- James, M. (1995). *The promise of crime prevention: The Elmira Prenatal/Early Infancy Project*. Canberra: Australian Institute of Criminology.
- Jessor, R. (1992). Risk behaviour in adolescence: A psychosocial framework for understanding and action. In D. E. Rogers & E. Ginzburg (Eds.), *Adolescents at risk: Medical and social perspectives*. Boulder, CO: Westview Press.
- Jessor, R. (1993). Successful adolescent development among youth in high-risk settings. *American Psychologist*, 48(2), 117-126.
- Johannessen, P., Bandara, J. S., & Smith, C. (2004). An analysis of the Sri Lankan conflict using an economic tool. *Economic Analysis and Policy*, 33(2), 163-188.
- Johannesson, M., Pliskin, J. S., & Weinstein, M. C. (1993). Are healthy years equivalent to an improvement over quality-adjusted life years? *Medical Decision Making*, 13, 281-286.
- Johnson, D. L. (2006a). Parent-Child Development Center follow-up project: Child Behavior problem results. *The Journal of Primary Prevention*, 27(4), 391-407.
- Johnson, D. L. (2006b). Parent-child development center follow-up project: Child behaviour problem results. *The Journal of Primary Prevention*, 27(4), 391-407.
- Johnson, D. L., & Blumenthal, J. (2004a). The Parent Child Development Centres and school achievement. *The Journal of Primary Prevention*, 25(2), 195-209.

- Johnson, D. L., & Blumenthal, J. B. (2004b). The parent-child development centres and school achievement: A follow-up. *The Journal of Primary Prevention*, 25(2), 195-209.
- Juniper, E. F., Guyatt, G. H., & Jaeschke, R. (1996). How to develop and validate a new health-related quality of life instrument. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 49-58). Philadelphia, New York: Lippincott-Raven.
- Kaplan, R. M., & Anderson, J. P. (1996). The general health policy model: An integrated approach. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd ed., pp. 309-322). Philadelphia: Lippincott-Raven.
- Kaplan, R. M., Anderson, J. P., Wu, A. W., Mathews, C., Kozin, F., & Orenstein, D. (1989). The Quality of well-being scale. *Medical Care*, 27(3), s27-s43.
- Kaplan, R. M., Bush, J. W., & Berry, C. C. (1976). Health status: Types of validity and the index of well-being. *Health Services Research* 11, 11(4), 478-507.
- Karoly, L. A., Greenwood, P., Everingham, S. S., Hoube, J., Kilburn, M. R., Rydell, C. P., Sanders, M., & Chiesa, J. (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions*. Santa Monica, CA: RAND.
- Keating, D. P., & Hertzman, C. (1999). Modernity's paradox. In D. P. Keating & C. Hertzman (Eds.), *Developmental health and the wealth of nations: Social, biological and educational dynamics* (pp. 1-18). New York: The Guilford Press.
- Keeney, R., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value trade-offs*. New York: Wiley.

- Klaus, P. (1994). *The costs of crime to victims*: U.S Department of Justice.
- Kleindorfer, P., Kunreuther, H., & Schoemaker, P. (1993). *Decision sciences - An integrative perspective*. New York: Cambridge University Press.
- Krantz, D. H. (1972). A theory of magnitude estimation and cross-modality matching. *Journal of Mathematical Psychology*, 9(2), 168-199.
- Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R., & Heintzelman, M. (2002). Age, health, and willingness to pay for mortality risk reductions: A contingent valuation survey of Ontario residents. *The Journal of Risk and Uncertainty*, 24(2), 161-186.
- Lally, R. J., Mangione, P. L., & 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 D. R. Powell (Ed.), *Parent education as early childhood intervention: Emerging directions in theory, research and practice* (Vol. 3, pp. 79-104). New Jersey: Ablex Publishing Corporation.
- Laub, J. H., & Sampson, R. J. (2003). *Shared beginnings, divergent lives: Delinquent boys to age 70*. Massachusetts: Harvard University Press.
- Laub, J. H., & Sampson, R. J. (2005). A general age-graded theory of crime: Lessons learned and the future of life-course criminology. In D. P. Farrington (Ed.), *Advances in criminological theory, Volume 13: Testing integrated developmental/life course theories of offending*. New Brunswick, NJ: Transaction Publishers.
- Lazar, I., & 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(2-3), 1-151.

- Lazar, I., Darlington, R., Murray, H., Royce, J., Snipper, A., & Ramey, C. T. (1982). Lasting effects of early education: A report from the Consortium for Longitudinal Studies. *Monographs of the Society for Research in Child Development, 47*(2-3), 1-151.
- Lee, V. E., Schnur, E., & Brooks-Gunn, J. (1988). Does Head Start work? A 1 - year follow-up comparison of disadvantaged children attending HEad Start, no preschool, and other preschool programs. *Developmental psychology, 24*(2), 210-222.
- Levenstein, P., Levenstein, S., Shiminski, J. A., & 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*(2), 267-285.
- Leventhal, J. M. (2001). The prevention of child abuse and neglect: Successfully out of the blocks. *Child Abuse and Neglect, 25*(2), 431-439.
- Levin, H. M., & McEwan, P. J. (2001). *Cost-effectiveness analysis* (2nd ed.). London: Sage Publications.
- Lipsey, M. W., & Wilson, D. B. (2000). *Practical meta-analysis* (Vol. 49). Thousand Oaks, CA: SAGE Publications.
- Little, I. M. D. (1950). *A critique of welfare economics*. Oxford: Oxford University Press.
- Little, M. (1999). Prevention and early intervention with children in need: Definitions, principles and examples of good practice. *Children and Society, 13*, 304-316.
- Love, J. M., Kisker, E. E., Ross, C., Constantine, J., Boller, K., Chazan-Cohen, R., Brady-Smith, C., Fuligni, A. S., Raikes, H., Brooks-Gunn, J., Tarullo, L. B., Schochet, P. Z., Paulsell, D., & Vogel, C. (2005). The effectiveness of Early

- Head Start for 3-year-old children and their parents: Lessons for policy and programs. *Developmental Psychology*, 41(6), 885-901.
- Luce, R. D., & Raiffa, H. (1957). *Games and decisions*. New York: John Wiley and Sons, Inc.
- MacLeod, J., & Nelson, G. (2000). Programs for the promotion of family wellness and the prevention of child maltreatment: A meta-analytic review. *Child Abuse and Neglect*, 24(9), 1127-1149.
- MacMillan, H. L., MacMillan, J. H., Offord, D. R., Griffith, L., & MacMillan, A. (1994). Primary prevention of child physical abuse and neglect: A critical review. *Journal of Child Psychology and Psychiatry*, 35(6), 835-856.
- Mandelblatt, J. S., Fryback, D. G., Weinstein, M. C., Russell, L. B., Gold, M. R., & Hadorn, D. C. (1996). Assessing the effectiveness of health interventions. In M. R. Gold, J. E. Siegel, L. B. Russell & M. C. Weinstein (Eds.), *Cost-Effectiveness in health and medicine* (pp. 135-175). New York: Oxford University Press.
- Mankiw, G. (2001). *Principles of economics* (2nd ed.). Sydney: Harcourt College Publishers.
- Manning, M. (2004). *Measuring the costs of community-based developmental prevention programs in Australia*. Unpublished Masters (Hons), Griffith University, Brisbane.
- Manning, M., Homel, R., & Smith, C. (2006). Economic evaluation of a community-based early intervention program implemented in a disadvantaged urban area of Queensland. *Economic Analysis and Policy*, 36(1 & 2), 99-120.



- Marshall, J., & Watt, P. (1999). *Child behaviour problems: A literature review of the size and nature of the problem and prevention interventions in childhood*. Perth, Western Australia: The Interagency Committee on Children's Futures.
- Maxfield, M. G., & Babbie, E. (2001). *Research methods for criminal justice and criminology*. Belmont, CA: Wadsworth.
- Mayhew, P. (2003). Counting the costs of crime in Australia. *Australian Institute of Criminology: Trends and issues in crime and criminal justice*, 247.
- McCain, M. N., & Mustard, F. J. (1999). *Early years study: Reversing the real brain drain*. Toronto: Publications Ontario.
- McCain, M. N., & Mustard, F. J. (2002). *The Early Years Study three years later: From early child development to human development: Enabling communities*. Toronto, Ontario: The Founders' Network of the Canadian Institute for Advanced Research.
- McCain, M. N., Mustard, F. J., Coffey, C., Gordon, M., Comis, J., Offord, D., Desjardins, J., Sullivan, T., Ferron, R., Will, C., Minz Geneen, F., & Williams, R. (1999). *Early Years Study: Reversing the real brain - Final Report*. Ontario: Ontario Children's Secretariat.
- McCarthy, D. A. (1972). *Manual for the McCarthy scales of children's abilities*. New York: Psychological Corporation.
- McConachie, H., & Waring, M. (2002). Parental coping and support. In I. Sclare (Ed.), *Child psychology portfolio*. Brisbane: NFER-Nelson.
- McLoyd, V. C. (1998a). Children in poverty: Development, public policy, and practice. In W. Damon, I. E. Sigel & K. A. Renninger (Eds.), *Handbook of child psychology* (Vol. 4: Child psychology in practice, pp. 135-208). New York: John Wiley and Sons.

- McLoyd, V. C. (1998b). Children in poverty: Development, public policy, and practice. In W. Damon, I. E. Sigel & K. A. Renniger (Eds.), *Handbook of child psychology* (Vol. 4: Child psychology in practice). New York: John Wiley and Sons.
- Medley, D. M., & Klein, A. A. (1957). Measuring classroom behaviour with a pupil-reaction inventory. *The Elementary School Journal*, 57(6), 315-319.
- Mehrez, A., & Gafni, A. (1989). Quality-adjusted life years, utility theory, and healthy-years equivalents. *Medical Decision Making*, 9, 142-149.
- Mehrez, A., & Gafni, A. (1993). Healthy-years equivalents versus quality-adjusted life years: In pursuit of progress. *Medical Decision Making*, 13, 287-292.
- Meisels, S., & Shonkoff, J. (2000). Early childhood intervention: A continuing evolution. In S. Meisels & J. Shonkoff (Eds.), *Handbook of Early Childhood Intervention* (pp. 3-31). New York: Cambridge University Press.
- Meyer, L. A. (1984). Long-term academic effects of the Direct Instruction Project Follow Through. *The Elementary School Journal*, 84(4), 380-394.
- Miller, L., & Bizzell, R. P. (1983). Long-term effects of four preschool programs: Sixth, seventh, and eighth grades. *Child Development*, 54(3), 727-741.
- Mongin, P. (1997). Expected utility theory. In J. Davis, W. Hands & U. Maki (Eds.), *Handbook of economic methodology* (pp. 342-350). London: Edward Elgar.
- Mrazek, P. J., & Brown, C. H. (2002). *Final report: An evidence-based literature review regarding outcomes in psychosocial prevention and early intervention in young children*. Toronto: Invest in Kids Foundation.
- Murray, C. J. L. (1996). Rethinking DALYs. In C. J. L. Murray & A. D. Lopez (Eds.), *The global burden of disease: Global burden of disease and injury series* (pp. 1-98). Cambridge: Harvard University Press.

- Murray, C. J. L., Salomon, J. A., Mathers, C. D., & Lopez, A. D. (Eds.). (2002). *Summary measures of population health: Concepts, ethics, measurement and applications*. Geneva: World Health Organisation.
- Nagin, D. (2001). Measuring the economic benefits of developmental prevention programs. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 28, pp. 347-384). Chicago: The University of Chicago Press.
- Nease, R. F. (1996). Do violations of the axioms of expected utility theory threaten decision analysis? *Medical Decision-Making*, 16(4), 399-403.
- Nelson, G., Laurendeau, M., Chamberland, C., & 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 I. Prilleltensky, G. Nelson & L. Peirson (Eds.), *Promoting family wellness and preventing child maltreatment: Fundamentals for thinking and action* (pp. 220-272). Toronto: University of Toronto Press.
- Nelson, G., Westhues, A., Laurier, W., & MacLeod, J. (2003). A meta-analysis of longitudinal research on preschool prevention programs for children. *Prevention and Treatment*, 6, 1-35.
- O'Brien Caughy, B., DiPietro, J. A., & Strobino, D. M. (1994). Day-care participation as a protective factor in the cognitive development of low-income families. *Child Development*, 65(2), 457-471.
- Office for Research Griffith University. (2007). *Research ethics manual: Booklet 22 - Informed consent in human research*, 2007, from <http://www.griffith.edu.au/or/ethics/humans/>
- Olds, D. (2002). Prenatal and infancy home visiting by nurses: From randomized trials to community replication. *Prevention Science*, 3(3), 153-172.

- Olds, D., Eckenrode, J., Henderson, C. R., Kitzman, H., Powers, J., Cole, R., Sidora, K., Morris, P., Pettitt, L., & Luckey, D. (1997). Long-term effects of home visitation on maternal life course and child abuse and neglect: Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association*, 278(8), 637-643.
- Olds, D., Henderson, C. R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidora, K., Morris, P., & Powers, J. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behaviour. *Journal of the American Medical Association*, 280(14), 1238-1244.
- Olds, D., Henderson, C. R., & 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(1), 89-98.
- Olds, D., Henderson, C. R., Phelps, C., Kitzman, H., & Hanks, C. (1993). Effects of prenatal and infancy nurse home visitation on government spending. *Medical Care*, 31, 155-174.
- Olds, D., & Korfmacher, J. (1998). Findings from a program of research on prenatal and early childhood home visitation: Special issue introduction. *Journal of Community Psychology*, 26(1), 1-3.
- Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, 8(2), 157-159.
- Parks, G. (2000). *The High/Scope Perry Preschool Project*. Washington: U.S. Department of Justice.
- Patterson, G., Reid, J., & Dishion, T. (1992). *Antisocial boys*. Eugene, OR: Castalia Press.

- Pavelski, P. R., Greif, J., Bates, M., & Furlong, M. (2001). *Santa Barbara School Readiness Scale: School readiness needs of Latino preschoolers in Santa Barbara County*. Santa Barbara: Centre for School-Based Youth Development: Gevirtz Graduate School of Education, University of California.
- Piotrkowski, C. S., Botsko, M., & Matthews, E. (2000). Parents' and teachers' beliefs about childrens school readiness in a high-need community. *Early Childhood Research Quarterly, 15*(4), 537-558.
- Pollard, J., Hawkins, J., & Arthur, M. (1999). Risk and protection: Are both necessary to understand diverse behavioural outcomes in adolescence? *Social Work Research, 23*(3), 145-158.
- Queensland Council of Social Services. (1999). *People & places - A profile of growing disadvantage in Queensland*. Brisbane: Queensland Council of Social Service Inc. and Social Action Office.
- Rachman, S. J., & Wilson, G. T. (1980). *The effects of psychological therapy* (2nd ed.). Oxford: Pergaman.
- Ramey, C. T., & Landesman Ramey, S. (1998). Early intervention and early experience. *American Psychologist, 53*(2), 109-120.
- Ramey, S. L., & Ramey, C. T. (1998). The transition to school: Opportunities and challanges for children, families, educators, and communities. *The Elementary School Journal, 98*(4), 293.
- Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for children at-risk. *Developmental Psychology, 30*(6), 787-804.
- Reynolds, A. J., Ou, S.-R., & 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(5), 1299-1328.
- Reynolds, A. J., Temple, J. A., Robertson, D. L., & 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(18).
- Richardson, J. (1990). *Economic assessment of health care: Theory and practice*: Monash University, Centre for Health Program Evaluation.
- Richardson, J., Olsen, J. A., Hawthorne, G., Mortimer, D., & Smith, R. (1998). *The measurement and valuation of quality of life in economic evaluation*. West Heidelberg, Victoria: The Centre for Health Program Evaluation.
- Richman, N., & Graham, P. J. (1971). A behavioural screening questionnaire for use with three-year-old children. *Journal of Child Psychology and Psychiatry*, 12(1), 5-33.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological bulletin*, 86, 638-641.
- Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 231-244). New York: Russell Sage Foundation.
- Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, 118(2), 183-192.
- Russell, L. B., Siegel, J. E., Daniels, N., Gold, M. R., Luce, B. R., & Mandelblatt, J. S. (2001). Cost-effectiveness analysis as a guide to resource allocation in

- health: Roles and limitations. In M. R. Gold, J. E. Siegel, L. B. Russell & M. C. Weinstein (Eds.), *Cost-effectiveness in health and medicine* (pp. 1-21). New York: Oxford University Press.
- Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: Preliminary findings. *Journal of Child Psychology and Psychiatry*, 8, 1-11.
- Rutter, M. (2007). Sure Start local programmes: an outsider's perspective. In J. Belsky, J. Barnes & E. Melhuish (Eds.), *The national evaluation of Sure Start: Does area-based early intervention work?* (pp. 197-210). Bristol: The Policy Press: The University of Bristol.
- Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15, 234-281.
- Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.
- Saaty, T. L. (1986). Axiomatic foundation of the analytic hierarchy process. *Management Science*, 32(7), 841-855.
- Saaty, T. L. (1990). *Multicriteria decision making: The Analytical Hierarchy Process*. Pittsburgh, USA: RWS Publications.
- Saaty, T. L. (1994). "How to make a decision: the analytic hierarchy process". *Interfaces*, 24(6), 19-43.
- Saaty, T. L. (2000). *Fundamentals of decision making and priority theory* (2nd ed.). Pittsburgh, USA: RWS Publications.
- Saaty, T. L. (2001). *Decision making for leaders* (3rd ed.). Pittsburgh, PA: RWS Publications.
- Saaty, T. L., & Vargas, L. G. (1984). The legitimacy of rank reversal. *Omega*, 12, 513-516.

- Sanders, M. (1999). Triple P-Positive Parenting Program: Towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clinical Child and Family Psychology Review*, 2(2), 71-90.
- 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(3), 1-17.
- Sanders, M., Markie-Dadds, C., & 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*.
- Schoemaker, P. (1992). Subjective expected utility theory revisited: A reductio ad absurdum paradox. *Theory and Decision*, 33(1), 1-21.
- 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., & Weikart, D. P. (1993). *Significant benefits: The High/Scope Perry Preschool Study through age 27*. Ypsilanti, MI: High/Scope Press.
- Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, S. W., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry Preschool study through age 40*. Ypsilanti, MI: High/Scope Press.
- Sen, A. (1991). Utility: Ideas and terminology. *Economics and Philosophy*, 7, 277-283.



- Sen, A. (1995). Rationality and social choice. *American Economic Review*, 85(1), 1-24.
- Sen, A. , Nussbaum, M. (1996). Functioning and capability: The foundation of Sen's and Nussbaum's development ethic. *Political Theory* 24 (4), 584-612.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review*, 17(8), 881-901.
- Smith, M., & Glass, G. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32(7), 752-760.
- Solso, R. L. (1997). *Mind and brain sciences in the 21st Century*. Cambridge, Mass: MIT Press.
- Spilker, B. (1996a). Evaluating and comparing life events data with quality of life data. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd Edition ed., pp. 467-475). Philadelphia: Lippincott-Raven Publishers.
- Spilker, B. (1996b). Introduction. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (2nd ed., pp. 1-10). Philadelphia, New York: Lippincott-Raven Publishers.
- Spilker, B., & Revicki, D. A. (1996). Taxonomy of Quality of Life. In B. Spilker (Ed.), *Quality of life and pharmacoeconomics in clinical trials* (pp. 25-40). Philadelphia, New York: Lippincott\_Raven Publishers.
- Sprigle, J. E., & Schaefer, L. (1985). Longitudinal evaluation of the effects of two compensatory preschool programs on fourth- through sixth-grade students. *Developmental Psychology*, 21(4), 702-708.

- Stanley, F., Richardson, S., & Prior, M. (2005). *Children of the lucky country? How Australia has turned its back on children and why children matter*. Sydney: Pan Macmillan.
- Takanishi, R. (1993). The opportunities of adolescence-Research, interventions, and policy. *American Psychologist*, 48(2), 85-87.
- Tarbell, D. S., & Saaty, T. L. (1980). The conflict in South Africa. *Journal of Peace Science*, 4, 151-168.
- The Developmental Crime Prevention Consortium. (1999). *Pathways to prevention: Developmental and early intervention approaches to crime in Australia*. Brisbane: National Crime Prevention: Commonwealth Attorney-General's Department,.
- Torrance, G. W. (1976). Social preferences for health states: An empirical evaluation of three measurement techniques. *Socio-Economic Planning Sciences*, 10(3), 129-136.
- Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics*, 5(1), 1-30.
- Torrance, G. W., Boyle, M. H., & Horwood, S. P. (1982). Application of multi-attribute utility theory to measure social preferences for health states. *Operations Research*, 30(6), 1043-1069.
- Torrance, G. W., Feeny, D. H., Furlong, W. J., Barr, R. D., Zhang, Y., & Wang, Q. (1996a). Multiattribute utility function for a comprehensive health status classification system: Health Utilities Index Mark 2. *Medical Care*, 34(7), 702-722.

- Torrance, G. W., Furlong, M., Barr, R. D., Zhang, Y., & Wang, Q. (1996b). Multi-attribute utility function for a comprehensive health status classification system: Health utilities index mark 2. *Medical Care*, 34(7), 702-722.
- Torrance, G. W., Furlong, M., Feeny, D. H., & Boyle, M. H. (1995). Multi-attribute preference functions. *Pharmaco Economics*, 7, 503-520.
- Torrance, G. W., Siegel, J. E., & Luce, B. R. (1996c). Framing and designing the cost-effectiveness analysis. In M. R. Gold, J. E. Siegel, L. B. Russell & M. C. Weinstein (Eds.), *Cost-effectiveness in health and medicine* (pp. 54-81). New York: Oxford University Press.
- Tremblay, R. E., Masse, B., Perron, D., LeBlanc, M., Schwartzman, A. E., & Ledingham, J. E. (1992). Early disruptive behavior, poor school achievement, delinquent behavior, and delinquent personality: Longitudinal analyses. *Journal of Consulting and Clinical Psychology*, 60(1), 64-72.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297-323.
- Udo, G. G. (2000). Using analytical hierarchy process to analyze the information technology outsourcing decision. *Industrial Management & Data Systems*, 100(9), 421-429.
- Vargas, L. G. (1984). Reciprocal matrices with random coefficients. *Journal of Mathematical Modeling*, 3, 69-81.
- Viscusi, K. (1993). The value of risks to life and health. *Journal of economic literature*, 31(4), 1912-1946.
- Viscusi, K. (2000). The Value of life in legal contexts: Survey and critique. *American Law and Economics Association*, 2(1), 195-222.

- Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behaviour*. Princeton, NJ: Princeton University Press.
- Von Winterfield, D., & Edwards, W. (1986). *Decision analysis and behavioural research*. Cambridge, U.K: Cambridge University Press.
- Wagner, M. M., & Clayton, S. L. (1999). The Parents as Teachers program: Results from two demonstrations. *The Future of Children*, 9(1), 91-115.
- Walker, J. (1996). *Estimates of the costs of crime in Australia*. Canberra: Australian Institute of Criminology.
- Wang, Y.-M., & Elhag, T. M. S. (2006). An approach to avoiding rank reversal in AHP. *Decision Support Systems*, 42, 1474-1480.
- Watson, S. R., & Freeling, N. S. (1982). Assessing attribute weights. *Omega*, 10(6), 582-585.
- Weikart, D. P., Bond, J. T., & McNeil, J. T. (1978). *The Ypsilanti Perry Preschool Project: Preschool years and longitudinal results through fourth grade*. Ypsilanti, MI: High/Scope Educational Research Foundation.
- Weikart, D. P., & Schweinhart, L. J. (1992). High/Scope preschool program outcomes. In J. McCord & R. E. Tremblay (Eds.), *Preventing antisocial behaviour: Interventions from birth through adolescence* (pp. 67-86). New York: The Guilford Press.
- Weissberg, R. P., & Greenberg, M. T. (1998). School and community competence-enhancement and prevention programs. In W. Damon, I. E. Sigel & K. A. Renniger (Eds.), *Handbook of child psychology* (Vol. 4). New York: Wiley.
- Welsh, B. C. (2001). Economic costs and benefits of early developmental prevention. In R. Loeber & D. P. Farrington (Eds.), *Child delinquents: Development*,

- intervention, and service needs* (pp. 339-358). Thousand Oaks, California.: Sage Publications.
- Wilkinson, I. (2002). Families and relationships. In I. Sclare (Ed.), *Child psychology portfolio*. Brisbane: NFER-Nelson.
- Williams, A. (1993). Cost-benefit analysis: Applied welfare economics or general decision aid. In A. Williams & E. Giardina (Eds.), *Efficiency in the public sector*. London: Edward Elgar.
- Wolf, F. M. (1986). *Meta-Analysis: Quantitative methods for research synthesis*. Newbury Park, California: Sage Publications.
- World Health Organisation. (1986). *Young people's health - a challenge for society: Report of a study group on young people and health for all by the year 2000* (No. Technical Report Series, No. 731). Geneva: World Health Organisation.
- World Health Organisation. (1996). *Introducing the WHOQOL instruments*. Retrieved 2006, from <http://www.who.int/evidence/assessment-instruments/qol/ql1.htm>
- World Health Organisation. (2007). *Disability adjusted life years*. Retrieved 2007, from <http://www.who.int/healthinfo/boddaly/en/print.html>
- Wortman, P. M. (1994). Judging research quality. In H. Cooper & L. Hedges (Eds.), *The handbook of research synthesis* (pp. 97-110). New York: Russell Sage Foundation.
- Xiang, Z., & Schweinhart, L. J. (2002). *Effects five years later: The Michigan School Readiness Program evaluation through age 10*. Michigan: High/Scope Educational Research Foundation.
- Xiang, Z., Schweinhart, L. J., Hohmann, C., Smith, C., Storer, E., & Oden, S. (2000). *Points of light: Third year report of the Michigan School Readiness evaluation*. Michigan: High/Scope Educational Research Foundation.

Yang, J., & Lee, H. (1997). An AHP decision model for facility location selection.

*Facilities*, 15(9/10), 241-254.

Yoshikawa, H. (1994). Prevention as cumulative protection: Effects of early family

support and education on chronic delinquency and its risks. *Psychological*

*bulletin*, 115(1), 28-54.

Zigler, E., Taussig, C., & Black, K. (1992). Early childhood intervention: A

promising preventative for juvenile delinquency. *American Psychologist*,

47(8), 997-1006.