

Meeting the Behaviour Challenge: Effective Groupwork for Schools

PhD Thesis

by

Toby Quibell

School of Education, Communication and Language Sciences

University of Newcastle upon Tyne

NEWCASTLE UNIVERSITY LIBRARY

204 26688 X

Thesis L8094

Abstract	5
Chapter One: Literature	7
Scenarios from the Real World	7
Introduction	8
Section 1.1: The Challenging Context	10
Section 1.2: Coping with the Scale of Disaffection	16
1.2.1: Group Intervention Approaches	16
1.2.2: Targeted Prevention Approaches	18
1.2.3: Non-Clinic Real World Approaches	19
Section 1.3: Coping with the Complexity of Disaffection	22
Section 1.4: Coping with the Need to Demonstrate Outcome	30
Section 1.5: Analysing the Evidence of Research Programmes	34
1.5.1: the prevention component	36
1.5.2: the non-clinic setting and group intervention component	39
1.5.3: The Complexity Component	42
1.5.4: Key studies	47
Section 1.6: Trigger for AGI Research Project	55
Section 1.7: Definition of AGI Research Project	58
Section 1.8: Importance of AGI Research Project	60
Reflective Commentary	60
Chapter Two: Questions & Hypotheses	63
Introduction	63
Section 2.1: Questions	64
Section 2.2: Summary of Design	66
Section 2.3: From Questions to Hypotheses	67
Section 2.4: Hypotheses	73
Hypothesis 1	73
Hypothesis 1 (sub)	74
Hypothesis 2	74
Hypothesis 3	75
Hypothesis 4	75
Hypothesis 5	76
Hypothesis 5(sub)	77
Hypothesis 6	77
Hypothesis 7	78
Hypothesis 8	78
Hypothesis 9:	79
Reflective Commentary	79
Chapter Three: Intervention	81
Introduction	81
Section 3.1: The Curriculum Study Group (CSG)	84
Section 3.2: The Action GroupSkills Intervention	89
3.2.1: Procedures for Setting up a Therapeutic Group	89
3.2.2: Clear Aims, Boundaries and Behaviour Outcomes	91
3.2.3: Activities to support Aims and Outcomes	96
3.2.4: Therapeutic Models	101
Section 3.3: Reflective Commentary	103
Section 3.4: Selected Lesson Plans	105
3.4.1: Lesson Plan for Week One	105
3.4.2: Lesson Plan for Week Five	109

3.4.3: Lesson Plan for Week Nine.....	114
Chapter Four: Methodology.....	117
Introduction.....	117
Section 4.1: Research Design:	118
4.1.1: Relating conceptual frameworks to AGI Design.....	118
4.1.2: Theory in the Bottom-up Conceptual Framework.....	131
4.1.3: The Concerns of AGI Methodology.....	134
4.1.4: Choice of Methodology and Justification	141
4.1.5: Summary of Methodological Choices	146
Section 4.2: Subjects, Instruments and Procedures	148
4.2.1: Subjects.....	148
4.2.2: Data Collection Instruments	155
4.2.3: Data Collection Procedures	167
Section 4.3: Data Analysis Procedures	171
4.3.1: Reporting Conventions.....	171
4.3.2: Analysis and Effect Size.....	172
4.3.3: Maximising Reliability	176
4.3.4: Hypothesis 9.....	179
Section 4.4: Applied Ethics	182
Section 4.5: Issues of Reflexivity and Bias.....	187
Reflective Commentary.....	189
Chapter Five: Results	191
Introduction.....	191
Hypotheses 1-4	191
Section 5.1: Hypothesis 1	192
Hypothesis 1a & 1(sub)a.....	194
Hypothesis 1b & 1(sub)b.....	198
Hypothesis 1c & 1(sub)c	205
Hypothesis 1d & 1(sub)d:	208
Section 5.2: Hypothesis 2	213
Hypothesis 2a.....	214
Hypothesis 2b.....	218
Hypothesis 2c	223
Total Problem T-scores	224
Hypothesis 2d.....	226
Section 5.3: Hypothesis 3	231
Hypothesis 3a.....	232
Hypothesis 3b.....	234
Hypothesis 3c	239
Hypothesis 3d.....	242
Section 5.4: Hypothesis 4	246
Hypothesis 4a.....	247
Hypothesis 4b.....	250
Hypothesis 4c	256
Hypothesis 4d.....	260
Introduction to Hypotheses 5 – 9	264
Section 5.5: Hypothesis 5	265
Hypotheses 5a & 5(sub)a.....	268
Hypothesis 5b & 5(sub)b	270
Hypotheses 5c & 5(sub)c.....	272
Section 5.6: Hypothesis 6	275

Hypothesis 6a	277
Hypothesis 6b	278
Hypothesis 6c	281
Section 5.7: Hypothesis 7	283
Hypothesis 7a	284
Hypothesis 7b	286
Hypothesis 7c	287
Section 5.8: Hypothesis 8	288
Hypothesis 8a	291
Hypothesis 8b	293
Hypothesis 8c:	295
Section 5.9: Hypothesis 9	297
Hypothesis 9a	299
Hypothesis 9b	300
Hypothesis 9c:	302
Chapter Six: Interpretation and Discussion of Findings	304
Introduction and Summary of Findings	304
Section 6.1: Interpretation of Hypotheses	307
6.1.1: Hypothesis 1	307
6.1.2: Hypothesis 2	310
6.1.3: Hypothesis 3	312
6.1.4: Hypothesis 4	314
6.1.5 Hypothesis 5	315
6.1.6 Hypothesis 6	316
6.1.7: Hypothesis 7	318
6.1.8: Hypothesis 8	319
6.1.9: Hypothesis 9	321
Section 6.2: Explanation and Synthesis	322
6.2.1: Explanation	323
6.2.2: Extraneous Effects Impacting on Data	330
6.2.3: Synthesis	338
Section 6.3. Discussion	340
6.3.1: This study in the context of other studies	340
6.3.2: Implications for Theory	345
Chapter Seven: Conclusion	350
Introduction	350
Section 7.1: Summary	350
Section 7.2: Potential Practical Impact	353
Section 7.3: Assessment of Success of Project	356
Section 7.4: Limitations of present Study	360
7.4.1: Limitations of Design	360
7.4.2: Limitations of Outcome	361
7.4.3: Limitations Arising from Bias	362
Section 7.5: Implications for Future Research	363
References	365
Appendix 1: The Dataset: Abbreviated Version	384
Appendix 2: Data Collection Instruments	385
Appendix 3: Direct Observation Protocol	386

Abstract

Youth disaffection is a widely used term that is used to denote a broad range of at-risk behaviours that affect large numbers of young people and have widely varying social and physiological causes. This intervention study examines the effectiveness of groupwork in meeting the challenges of scale and complexity of reducing the impact of disaffection in the school and home environments. Using a sample of 130 children from primary, middle and secondary schools in North Tyneside, two group interventions were run with randomised samples and data collected up to one-year post intervention. This study presents the findings from data gathered at four time points comparing a curriculum support intervention with a reflective therapeutic intervention. Both interventions were conducted with small groups of children withdrawn from class and measures were taken to control extraneous variables in a quasi-experimental design. Assessments were made at four time points through questionnaires administered to the children, their parents and their teachers. Questionnaire data were complemented with a direct observation protocol for measuring behaviours in the classroom. School attendance was also monitored. The effectiveness of groupwork was measured over time, and the interventions compared to each other using measures of statistical significance and effect size. The study found agreement between the children and their teachers that both forms of groupwork produced a reduction in problem behaviours and an increase in self-concept over the intervention period. This improvement was still detectable one-year post-intervention. The teachers and parents, but not the pupils, favourably compared the intervention with a three-month waiting list period where data were gathered in the absence of intervention. The teachers and parents distinguished between the intervention conditions over the intervention period, attributing marginally greater gains to the therapeutic intervention than to the curriculum intervention. At one year post-intervention, the teachers again distinguished between the conditions, attributing marginally greater gains to the therapeutic intervention than to the curriculum intervention. These findings were considered to support the use of groupwork in schools. In an age dominated by evidence-based approaches that look for specific interventions for categorised symptom-clusters, this study provides evidence from a rigorous methodology that clustering children according to teacher concern cutting across diagnostic categories can be the starting point for context-friendly interventions useful to those seeking community-based solutions to the complex social issue of disaffection.

Index of Tables

Table 1.5.4:	Match between Design Components of Key Studies and those Identified as Desirable	47
Table 3.2	Grid of Planned Week by Week Aims	88
Table 3.2.3a	Structure for Activities	97
Table 3.2.3b	Activity Skill Matrix	100
Table 4.2.1a:	Subjects	149
Table 4.2.1b:	Criteria for Inclusion in the GroupSkills Intervention Programme	151
Table 4.2.4:	Timetable of Intervention and Data collection	168
Table 4.3.2:	Comparing Cohen's and Hopkins' effect size categories	175
Table 4.3.4a	TRF Gain Scores T2-T4	179
Table 4.3.4a	YSR Gain Scores T2-T4	180
Table 4.3.4a	CBC Gain Scores T2-T4	180
Table 4.3.4d	School Scores According to Commitment Criteria (1-10)	180
Table 5.1a:	Descriptives & Analysis	195
Table 5.1b:	Descriptives & Analysis	199
Table 5.1c:	Descriptives & Analysis	205
Table 5.1d:	Descriptives & Analysis	209
Table 5.2a:	Descriptives & Analysis	215
Table 5.2b:	Descriptives & Analysis	218
Table 5.2c:	Descriptives & Analysis	224
Table 5.2d:	Descriptives & Analysis	226
Table 5.3a:	Descriptives & Analysis	232
Table 5.3b:	Descriptives & Analysis	235
Table 5.3c:	Descriptives & Analysis	240
Table 5.3d:	Descriptives & Analysis	243
Table 5.4a:	Descriptives, Analysis (ANOVA) & Analysis (Effect Size)	247
Table 5.4b:	Descriptives, Analysis (ANOVA) & Analysis (Effect Size)	250
Table 5.4c:	Descriptives, Analysis (ANOVA) & Analysis (Effect Size)	257
Table 5.4d:	Descriptives, Analysis (ANOVA) & Analysis (Effect Size)	261
Table 5.5:	Descriptives & Analysis	267
Table 5.6:	Descriptives & Analysis	276
Table 5.7:	Descriptives & Analysis	284
Table 5.8:	Descriptives & Analysis	290
Table 5.9	Correlation Between Rank Order of School Commitment and Gain Score (T2-T4)	298
Table 6.1a:	Null or Experimental Hypothesis Acceptance 1-4&9	305
Table 6.1b:	Null or Experimental Hypothesis Acceptance 4-8	306
Table 6.1c:	Effect sizes for Hypotheses 1-4	306
Table 6.1d:	Effect sizes for Hypotheses 5-8	306
Table 6.1e:	Attrition	307
Table 6.1.1	Post-Hoc Waiting List Analysis	309
Table 6.2.2a	Correlation coefficients between gain score and initial score and significance value (two tailed)	332
Table 6.2.2b	Attrition levels according to severity	334
Table 6.2.2c	Change over time: pattern of gains made by those with waiting list assessments (T1=1) and those without (T1=0)	336
Table 6.2.2d	Mean scores of those with waiting list assessments (T1=1) and those without (T1=0)	336
Table 6.3.1	Match between design components of key studies and those identified as desirable	341

Chapter One: Literature

Scenarios from the Real World

A classteacher finishes the introduction of a new topic to an unsettled class on a hot day. She asks the children to take out their workbooks and begin a new page by making a diagram of the flowchart they have developed together on the whiteboard. There is an immediate scuffle towards the back, and a head appears twisted with anger “Miss, he hit me, he fuckin hit me” The class teacher strides towards the back of the room, her finger hovering over the Bromcom alert button: already there are signs that the class will dissolve into chaos...

In the Child and Adolescent Mental Health clinic the waiting room is full. Michael Rowley is doing his best to sit still but both he and his mother are impatient. He is due another Ritalin tablet soon and she can feel his barely suppressed desire to jump up and kick the wall. The registrar opens a consulting room door and beckons them in, this clinic seems to have been going on forever and each review takes the best part of an hour. The Doctor hates herself for her own cynicism, but ADHD seems to have mushroomed following the realisation that there is a disability allowance that is attached to each diagnosis.

It has been a quiet day at ‘40-ODD’ the young people’s centre in the heart of the estate. The usual sad cases have turned up in the morning, desperate to avoid school, hiding from the wag-man, nonchalant of the consequence of their school absence, the truanting and the hide and seek is just part of the usual routine. The sad loners went off around lunchtime with Craig, and you knew what that meant – they would be on some disgusting sofa somewhere in a chemical haze, nothing too serious probably – tack or wobbly eggs. Later they might go out and have some fun: they know how to cause the maximum disruption while maintaining the maximum self-righteousness.

Introduction

In cities, towns and villages young people are involved in destructive and self-destructive behaviours, and it appears to be growing as a phenomenon. Whether it is literally destructive breaking, stealing or damaging property, or a self-destructive opting out of school by being a bit of a wag, a bit of a bully and living off the approval of the peer group, young people are impacting directly on those around them.

How do we go about changing the behaviour of these young people? Schools increasingly are open to the notion that they are not just places for curriculum, although teachers often wish to be let alone to teach their subject without having to manage the kind of behaviours that disfigure their lesson plans. CAMHS, educational welfare, educational psychology are only now emerging from a philosophy of treatment based on the individual consultation, and the communities in which these people congregate in large numbers continue to suffer the worst of crime and physical damage. The prisons are full of adults who have graduated from youth offending and the young offender institutions are having a crisis of confidence as to whether they do anything other than contain young people for the time they are present.

The resources seem to be unequal to the scale of the need. Received notions of fitting treatments and deterrents do not seem to be reducing the numbers involved in behaviours of concern. The root of these behaviours is variously characterised, but can be seen as complex mixture of failed attachments, poor role modelling, lack of academic ability and aspiration, pure bloody mindedness and scant positive life opportunities. Long-term solutions will have to engage with this complexity, for just as the behaviours are not uniform, neither are the causes.

The pressing question is whether an intervention can meet the challenges of scale (i.e. deal with a large number of people) while delivering content that engages with the complexity of the causes of behaviour. Knowing, moreover, that to produce more well-balanced criminals is no kind of outcome. Interventions need to impact upon the real world, affecting scenarios like those at the beginning of this chapter.

Can an intervention impact upon youth disaffection in such a way as to directly assist young people in developing positive futures, in such a way as to support the

development of safer communities and effective schooling, and in such a way so as to accommodate the demands on over-stretched public servants and services?

This piece of research sets out to answer the question of whether an evidence base can be built for an intervention that addresses these three concerns of scale, complexity and practical outcome. This piece of research is working towards solutions to a real world problem, by investigating group interventions in the school setting with youth at-risk of disaffection.

Glossary of terms

Youth: Children are here understood to be those aged between 0 and 18, but for the purposes of this research particular attention is given to those aged between 6 and 15. These children will be collectively referred to as 'youth'.

Disaffection: When children fall repeatedly short of received expectations with consequences that are damaging to their prospects, peer relationships and authority relationships the children will be spoken of as 'disaffected' or 'at-risk of disaffection'. The term disaffection is favoured over alternatives such as maladjusted, deprived, troubled, challenging, anti-social and disturbed because it is considered to hold an understanding of the importance of emotion, or affect, in the form of their disaffection.

Disturbance: When children's disaffection is of a severity to register on crime statistics and levels of clinical diagnosis it is termed disturbance. This is to differentiate severity of behaviours in the at-risk population and the clinical population.

Internalising and Externalising Behaviours: These terms are used throughout the AGI study, and originate from subscales aggregated by Achenbach in the battery of assessment instruments he developed (Achenbach 1993). The internalising aggregate of behaviours includes subscales relating to depressive, withdrawn symptoms, and the externalising aggregate includes those subscales relating to aggressive, conduct symptoms. When combined these aggregates give a score for total problems on the Achenbach instruments, but the terms internalising and externalising are used throughout the text to characterise depressive and aggressive behaviours respectively.

Medical Terminology: The term disaffection does not however preclude the associations of meaning contained in the alternative terms, but just places different emphasis on potential remedial action. In engaging with the literature, all the above terms may be borrowed in order to make sense of the emerging evidence. Words in common usage in

medical discourse such as treatment, symptom, syndrome, clinical psychopathology etc. are freely used, within an awareness of the sociological critique of a medical model of disturbance.

Clinical Category: This piece of medical terminology refers to the population of children with T-scores of 60 (1 standard deviation) or above on measures of disturbance. The term clinical levels is used with the understanding that the decision to set the level at 1 standard deviation is an arbitrary one, as there is no qualitative difference between behaviours either side of, but close to, this cut-off. In addition, with correlations between different measures below 1 there can be confusion about the percentage of children in the clinical category, as different children score above 60 according to the measure used.

At-Risk: The at-risk population is understood to be those young people who over a five-year period are likely to fall into a clinical category.

The AGI study: This research project investigates the effect of group interventions: a curriculum studies group (CSG) and a group using a therapeutic intervention named Action GroupSkills Intervention (AGI). For the sake of brevity, the whole research project shall be referred to as the AGI Study.

The Learning Challenge (TLC): TLC is the organisation founded by Toby Quibell to deliver groupwork programmes to at-risk children in schools. The AGI study was supervised by a research team, but where referring to aims and vision, it is the vision of TLC that is detailed.

Section 1.1: The Challenging Context

The AGI study is directed towards building an evidence base for an intervention that addresses the concerns of scale, complexity and practical outcome in youth disaffection. This section examines the literature in order to establish how widespread behaviours of disaffection are. Disaffected behaviours will include, under the definition advanced above, behaviours that do not register in many official figures. Arriving at a notion of the prevalence of disaffection will be done by beginning with the incidence of youth disturbance (that is severe and/or clinical behaviours) to get a fixed point from which to work.

Severe destructive behaviour exists in a number of forms that register on official statistics: youth crime statistics is one, prevalence of mental health syndromes is another. Schools provide data on achievement and attendance, but are focussed on demonstrating performance rather than looking at incidence of disturbance. None of the figures that emerge from the sources above captures the disaffection that underlies the severe behaviours that make it onto the published lists.

In describing the context for the AGI study, there is therefore the difficulty of official published statistics reporting only extreme examples of disaffected behaviour. Using crime as a starting point, there seems little doubt that offending rates have been on a steady rise over the last half of the twentieth century, with few national exceptions. Recent reviews (e.g. Youth Justice Board UK 2003; Farrington 1996; Smith 1995) confirm the picture. The reasons for this rise are elusive and complicated (Rutter, Giller, & Hagell 1998), but it does seem certain that juvenile levels of offending have risen over that same period given that a quarter of crime is committed by juveniles. In the United Kingdom (Farrington 1996), the figures for the numbers of juveniles found guilty or cautioned for indictable offences was at a level of 540,000 according to official statistics (Great Britain Home Office 2001).

Crime statistics are one way of getting at the hard figures about the prevalence of youth disturbance, but it is only part of the picture. The prevalence of childhood psychiatric disorder among young people in the United Kingdom is another set of 'hard' data that gives a glimpse of the disturbance felt by children.

The methodological problems in conducting prevalence studies of childhood psychiatric disorder mean that the prevalence reported by studies ranges from 1 per cent to 51 per cent (Roberts 1998), and the review conducted by Roberts took a mean level of prevalence at 15.8 per cent from 52 studies reported in 47 sources, with the United States and England being the most frequent sites for studies. Median rates were 8 per cent for preschoolers, 12 per cent for preadolescents, 15 per cent for adolescents and 18 per cent in studies including wider age ranges. A more recent single study surveyed mental health of the British child and adolescent population in 1999 and is distinguished by method and analysis (Ford, Goodman, & Meltzer 2003). This study looks at a representative sample of children and after collecting data through interview, matched

reported behaviours to the fourth edition of the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV 1994). The study found that at least one DSM-IV diagnosis was present in 9.5 per cent of the sample (Meltzer p1205). A fifth of these children (2.1 per cent) had 'not otherwise specified' diagnoses that failed to meet current DSM-IV criteria but were causing the child significant distress. This report put prevalence of clinical level of mental health need at a level of one in ten UK school children. The authors recognise that at 9.5 per cent the prevalence is at the lower end of the range reported by other studies, the most important reason for which being that the diagnostic criteria incorporated impact criteria designed to detect disorder that resulted in significant interference in the child's life and merited clinical intervention. An example of a recent survey conducted in the United Kingdom reports that 17% of parents regard their children as prone to stress, panic attacks and depression (Norwich Union Healthcare 2004). This study, commissioned as part of a national initiative to help parents get the best support for their children's emotional and physical wellbeing, questioned 400 parents of children aged 5 to 15, and while these figures should be regarded as indicative rather than authoritative, it helps to give a picture of the levels of concern and prevalence of mental health disorders.

Interestingly a parallel study in the United Kingdom (Meltzer et al. 2000) using the World Health Organisation's *International Classification of Diseases* (ICD-10 1993) found similar levels of mental disorder: 10.4 per cent of children in England Scotland and Wales were found to have some type of mental disorder. In a methodology that used random sampling of child benefit records to recruit and train interviewers for gathering data from parents and older children, this study diagnosed mental disorder not just on symptoms but on evidence of distress or interference with personal function. In three broad groupings, 5.3 per cent of children were found to have conduct disorders, 4.3 per cent emotional disorders and 1.4 per cent classed as hyperactive. The less common disorders such as autistic disorders, tics and eating disorders were attributed to half a per cent of the sampled population. Among 5 to 10 year olds, 10 per cent of boys and 6 per cent of girls had a mental disorder; among the 11 to 15 year olds the proportions were 3 per cent for boys and 10 per cent for girls.

The interest here is in the similarity with the study reporting DSM-IV disorders, giving some credence to the figure of 10 per cent prevalence of mental health disorder in UK

schoolchildren, and in the breakdown of disorder according to type and gender. Also importantly, a strong association was found separately between social class and family income and the mental health of the child. Children from families in social class V (unskilled occupations) were three times more likely to have a mental health problem than those from social class I (professionals). The prevalence for any disorder ranged from 16 per cent among children living in families with a gross weekly income of under £100, to 6 per cent in those families earning £500 per week or more. The pattern that emerges will surprise no one familiar with the issues of deprivation. In inner city communities where the problems of disaffection are acute, 16 per cent of children have a clinical level of mental health need that is of significant impairment to their function. In the average class of 30 children this equates to 5 children with significant impairment to their function, whether conduct, emotional or hyperactive.

These facts reveal a high prevalence of clinical levels in youth disturbance in classrooms and inner-city communities across the country that seems to gravitate around 10% of a mixed socio-economic population in the UK. This figure is given extra credence by Bird (1996) who summarized 13 major epidemiological studies of school age children and adolescents, all of whom refined their estimates of the prevalence of disorder by reporting the proportion of subjects whose disorder was of some clinical significance. This resulted in an average prevalence across 13 countries of 12 per cent. According to this reading of the literature, 1.2 in 10 children have a level of need requiring professional referral and specialised attention: it is altogether unsurprising that the strains on teachers, educational psychologists, psychiatrists and health workers is so great. These are children who will need accurate diagnosis, referral and treatment. What tends to happen is resource intensive, whether in one of the tertiary, or possibly quaternary, centres of child and adolescent mental health or through close monitoring in the school through the statementing process. Contact with the home environment during and post treatment is important and the treatment itself is likely to be clinic-based. As a result of the diagnosed need, it is likely that initial assessment is done individually and any further non-clinic treatment and monitoring is likely to be individual.

In an average primary school of 200 these considerations will apply to 32 children, in a high school of 800 these considerations will apply to 128 children. A local education authority might have 6 high schools each with 5 feeder primary schools, giving a back-

of-envelope total of 1800 children in a local authority with clinical level disturbance and all the detail of diagnosis treatment and management that it entails. In the North East of England there is one tertiary/quaternary centre for child and adolescent mental health that serves the region, covering 5 LEAs giving 9000 children to deal with. Most of those will be dealt with by local centres for child and adolescent mental health run by primary care teams, but even so the numbers present a huge challenge for services.

If the levels of disturbance are daunting at a figure of about 12%, what can be said of the levels of disaffection? The levels of sub-clinical and sub-criminal prevalence are more difficult to ascertain with any accuracy, but it is without doubt true that for every child with a clinical level of conduct disorder or emotional problem or involvement with the youth justice system, there are several with signs and symptoms of emotional/social impairment that have a probability of registering on the official figures in time. Getting at the levels of disaffection is a matter of extrapolation and logical deduction, but it is aided by evidence in the literature of the extent of the at-risk population.

For example, if account is taken of estimates that only 23 per cent of crime is detected (Great Britain Home Office 2001), it lends credibility to a more recent survey reporting almost half (48.5 per cent) of school aged UK children reported committing some kind of criminal offence in their lives (Beinart et al. 2002). This study reports around four out of ten 14 and 15 year olds in Year 10 said that they had 'ever' stolen or tried to steal something. Asked about the last year, a third of these students said they had vandalised property, but pains are made in the report to distinguish between participation in theft, vandalism and handling stolen goods and more serious property crimes, such as burglary, which were considerably less common (one in ten boys in year 11 said they had broken into a building to steal in the last year, including 4 per cent who reported doing so three or more times). Another report relating to the UK population states that "26% of mainstream young people [aged between 5 and 15] say that they have committed an offence in the last 12 months and 60 % of excluded young people say they have committed an offence in the last 12 months" (Youth Justice Board UK 2003).

These reports give an indication of the levels of at-risk behaviour in the youth population that does not necessarily register in official figures. If 26% of the mainstream has committed an offence this would indicate that in any one 12-month period, 26% could be

considered at-risk of criminal behaviour. The comparative evidence from mental health broadly supports this level of prevalence. So-called 'psychopathology in the community' epidemiological studies report levels of problems not exclusively at clinical levels. The Isle of Wight study puts prevalence at 6.8 per cent and an inner city London borough reported prevalence 25.4 per cent (Rutter et al. 1975). However, Bird points out that prior to 1980, studies differed markedly in the age groups included and methodologies (Bird 1996), and that the measures of psychopathology were neither based on diagnostic viewpoint nor empirically grounded. This makes it near to impossible to compare the Rutter study with (e.g.), the Ontario Child Health Study (Offord et al. 1987) reporting prevalence at 18.1 per cent. Overall, Bird found the studies' estimates of the prevalence of childhood psychopathology ranged from 6.6 per cent to 37 per cent with psychiatric disorders in these community surveys occurring in around 20 – 30 percent of school age children.

A more recent review (Roberts 1998) finds broad agreement about the difficulty in comparing studies before 1990 with those after, but draws particular attention to the effect on prevalence rates when using impairment criteria in determining "caseness". In common with the Ford study that gave a 9.5% prevalence after the use of impairment criteria (a significantly lower estimate than generally accepted), Roberts found that 'The prevalence rates adjusted for impairment were typically less (sometimes much less) than one half the prevalence rates based only on meeting symptom criteria.' (p721). Roberts also reports "With respect to case definition, DSM-III and DSM-III-R generated similar prevalence rates of 19%-23% and 20%-22% respectively, while clinical opinion yielded rates of 10%-14%." (p721). The striking co-incidence of the prevalence rates adjusted for impairment from the Ford study and those reported from clinical opinion suggests that clinicians are applying additional criteria from intuition or experience to moderate the prevalence that is derived from application of DSM criteria alone. Using impairment criteria and using intuition/experience reduce the prevalence levels by half.

In order to get a working figure for the prevalence of disaffection, the literature discussed above can be used to suggest a level of between 20% and 25%. The logic for this is that when DSM criteria are moderated by impairment criteria or experience, the resulting level of clinical prevalence is at between 10% and 13%. However the DSM criteria will have been applied on the basis of some evidence of symptoms to arrive at a figure of

20-22%, some of which will not be considered to be severe enough to be regarded as clinical. Using a level of prevalence for disaffection of between 20% and 25% takes the DSM diagnosis into account and also the figures from the crime statistics of 26% of mainstream population involved in offending behaviour. Any preventative intervention has to set its sights on 20-25% per cent of a youth population if it is to be effective in reducing the levels of disaffection that strangle our school system, our health and criminal justice system and choke the civic function of our inner cities.

Section 1.2: Coping with the Scale of Disaffection

These issues of scale and severity are a core concern when developing an intervention to deal with the symptoms of disaffection. Would the best response be to develop effective treatments to address clinical levels of disturbance when they reach the appropriate level? Or would it be better to intervene with a larger number when behaviours are less severe? Whatever conclusions are drawn from the literature about where to intervene in the scale of severity, the rise in prevalence of clinical need, and by inference sub-clinical behaviours, reported above makes the issue of matching treatment resource to incidence of behaviour, whatever its severity, a priority. Whether it is 10% of the child population in our clinics or 26% in our schools, there is a pressing need to be able to address the scale of need.

1.2.1: Group Intervention Approaches

Against this background, identifying and treating children with anti-social behaviour one-by-one has received much criticism. Firstly there is little evidence that treating severe anti-social behaviours is effective (Warren, Oppenheim, & Emde 1996), particularly in the case of early-onset conduct disorder (Hansen, Meissler, & Ovens 2000) and where effective services exist for severe problems, their complexity and attendant resourcing implications limit their wider application (Camodeca et al. 2003). Secondly, even if effective treatments were available, the number of children with mental health problems far exceeds the resources available for traditional mental health services in most communities. Estimates of the prevalence of clinical disorders in Canadian services outnumber the percentages of children using mental health services by factors of between 10 to 20 (Stilwell et al. 1997). Third, survey data indicate that the targeting of children's mental health services is far from accurate (Reynolds & Robertson 2003): it is

difficult to get children and families at highest need to access therapeutic services, even when these are available. For example there is evidence to suggest that many parents of high-risk children do not enrol in parent training programmes even when widely available (Webster-Stratton, Reid, & Hammond 2001). Finally the process of identification and referral takes time. During that interval, children not only suffer, but many will experience a downward spiral in function until behaviours are severe enough to warrant attention.

Group treatment approaches have always been present in the literature, but with the research emphasis remaining on the evidence-base for treatment modalities on specific diagnoses, (e.g. Fonagy et al. 2002), many research studies that do include adolescent groups include brief, cognitive behavioural treatment with specified diagnostic populations (e.g. Curry et al. 2003). This trend has created the impression, noted by Pollock & Kymissis (2001), that only certain types of technique are effective, since they have been the ones tested in research in recent years, and that the group setting is somehow incidental to the treatment modality. MacLennan (2000) amongst others, points out that with the advent of managed care, there is a pressure toward short-term and abbreviated treatment and a shortage of institutional funds to provide back-up support, but there is no reason why this should preclude the group situation as a context for intervention. At the same time there has been a proliferation of treatment models and of psychotropic drugs which can be used in conjunction with relationship and cognitive-based therapy but are more usually seen as an alternative (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002). This has not done much to raise the profile of group approaches as anything other than an addendum of setting in the rise of 'abbreviated treatment models'.

The research profile into the value of group factors (Yalom 1975) as effective in reducing the symptoms of disaffection is significantly lacking. It is difficult to find reliable data on the comparative effectiveness of different combinations of group and treatment approach (Azima 1996), and where research is done it rarely meets basic research criteria (Pollock & Kymissis 2001), consequently finding data on the range and extent of effective group practice is difficult (MacLennan 2000). However, where they do exist, studies on treatment models using the group context provide positive data (Lomonaco, Scheidlinger, & Aronsen 2000), although it is not clear whether this is an endorsement of

the treatment mode or the group setting. The Kolvin study is a little dated, although under scrutiny it holds up well (see below), but the very fact that it was conducted in the late 1970s and published in the early 1980s reveals a certain gap in the literature that this research hopes to address, because the common sense case for group intervention as an answer to the concerns of scale is strong.

1.2.2: Targeted Prevention Approaches

The acknowledged strain placed upon traditional child mental health services by the scale of need in youth disturbance has led researchers and practitioners to consider the possibility that resources might be more effectively spent in reducing the numbers of young people presenting at clinical level. Here again the common sense case is strong, but the data from preventative studies are hard to track down, complex to interpret and difficult to marshal.

Over the years, the results of group intervention programmes in producing superior cognitive skills, school adjustment, school attainment, favourable attitudes towards school, and better family and social adjustment have been both encouraging and disappointing (Kagitcibasi, Sunar, & Bekman 2001). Gains in cognitive ability have been regularly observed in the short term, but they dissipate in the longer term (Cicirelli 1969); on the other hand studies show lower drop-out rates, better school adjustment and other beneficial long term effects (Berrueta-Clement et al. 1984).

A problem with the early negative evaluations, mainly of the Headstart programme, was that they tended to consider only the first wave of results without waiting for later evidence regarding long-term effects. Later studies, looking at a wider range of outcome variables have revealed positive results pointing to sustained positive effects, especially better school adjustment and social integration, which are attributed mainly to motivational factors (Lazar & Darlington 1982). Currently there appears to be a general consensus on the long-term benefits of early enrichment, particularly for the social acceptance and adjustment of at-risk children (Campbell & Ramey 1994). There is also a recent focus on the possible long-term risk-prevention function of early childhood support programmes (Yoshikawa 1994). For example Yoshikawa points to three programmes, mostly involving centre-based early childhood education with parent training/family support, which show long-term reductions in antisocial or disaffected

behaviour. These are the Perry Preschool Project (Berrueta-Clement, Shweinhart, Barnett, & Epstein 1984), the Houston PCDC Program (Johnson & Walker 1987), and the Syracuse University Family Development research Program (Lally et al. 1988). Nevertheless, debate continues mostly with regard to the best approaches to be used such as direct (child focused) or indirect/mediated (parent-focused) approaches, the duration and scope of intervention and the type and purpose of service.

Careful consideration by clinicians, researchers and healthcare administrators of the development and evaluation of preventative initiatives (Durlak & Wells 1998) has led to a growth in acceptance of preventative approaches to meeting the challenge of anti-social behaviour. This is evidenced through theory (Yoshikawa 1994) and the positive outcomes of large scale preventative projects of the type detailed above. Prevention in the field of child and adolescent mental health has been classified into targeted and universal interventions (Offord 1987). The latter are aimed at the general population, or parts of the general population, regardless of whether they have a higher than average risk of developing a disorder and the majority of the studies mentioned fit into this category. However on the basis of the difficulty of reducing risk among those at low risk, the very large numbers required to demonstrate an effect on the incidence of new disorders, and their necessarily high cost, the applicability of universal interventions to the mental health field has been questioned. (Rotheram-Borus & Duan 2003). Targeted interventions avoid the exposure of large number of low risk individuals to an intervention and may be more cost effective. Targeted interventions that focus on a class of individuals who have been exposed to a risk factor have been termed selective interventions and aim to intervene when the behaviours exist in a sub-acute form for a population upon which the intervention will focus (secondary prevention). The fact that many dysfunctions continue across a lifespan underlies the importance of early intervention for children, both for a reduction in suffering at the time, and to attenuate or avoid impairment later on in life.

1.2.3: Non-Clinic Real World Approaches

Preventative treatment has shown success in clinical settings, as evidenced by the three programmes highlighted by Yoshikawa involving centre-based early education early childhood education (Yoshikawa 1994). However provision of this type is often affected by poor attendance and the fact that referral channels for children are adult dependent

(Fonagy, Moran, & Target 1993). The encouraging evidence from these controlled efficacy studies has led to recent attention turning to the examination of whether these treatments are effective in non-clinic 'real world' settings e.g. (Ginsburg & Drake 2002). So-called real-world settings are taken to mean the natural environments in which children gather, and are primarily taken to mean schools. Settings such as these go some considerable way to addressing the difficulty in conducting preventative activities in clinics or 'centers' with their associations of illness and treatment. In addition delivering interventions broadly under the heading of mental health services through the school system can address key financial and structural barriers that often prevent deprived communities from receiving much needed services (Kataoka et al. 2003).

Schools have long been identified as an ideal entry point for improving access to mental health services for children (Adelman & Taylor 2000) and are set to play a key role in the mental health of young children (The Children Act 2004) and their unique socialising potential has been recognised by researchers as well as teachers (Kolvin, Garside, Nicol, Macmillan, Wolstenholme, & Leitch 1981), (Brint, Contreras, & Matthews 2001), (Zsolnai 2002). Although there is wide recognition that school-based approaches to the management of delinquency have considerable potential, recent reviews of the literature typically conclude that 'studies to date have failed to demonstrate powerful positive effects' (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002), perhaps due to the perception that '...few programs have been rigorously evaluated in the real-world setting of schools' (Hoagwood 2000). However, school based interventions do overcome the disadvantage of clinic-based interventions (in that gains made in clinic are notoriously difficult to generalise) by seeking gains in the locus (or one of the main loci) of disturbance, and the school environment with its ecological setting combined with a relatively constant population, does provide a natural focus for different agencies affected by anti-social behaviour (Adelman & Taylor 2000). Recent thinking has focussed on the challenges of researching school-based prevention (Greenberg 2004), highlighting the need for integrated educational and therapeutic systems to build sustainability. Additional thinking has gone into the interface between prevention research and practice in schools (Kaftarian et al. 2004), giving cause for optimism that some of the concerns about the school as a place for effective intervention will be met in the near future.

Moreover school based interventions can be effective as shown in a major review of school-based programmes published in 1995 (Durlak 1995). The review draws attention to the effectiveness of secondary prevention programmes, and evaluates programmes targeting specific behaviours that include behavioural and social problems, substance use, prevention of academic problems and health education. The focus of school-based interventions have been social skills training and academic enrichment, with meta-analysis of social skills training studies indicating that children show social gains, at least in the short-term (Beelman, Pfingsten, & Losel 1994). However, studies also confirm intuitive expectations that interventions aimed at social skills and academic enrichment leave areas of anti-social behaviour unaffected (Lochman et al. 1993), resulting in small and sporadic effects attributable to intervention programmes (Hundert et al. 1999), more recently Durlak reported broad support for findings in his evaluation of the outcomes from 130 indicated preventative interventions (Durlak & Wells 1998) and Cowen demonstrated effects in a large study (Cowen et al. 1996).

Despite the tendency to focus on academic enrichment, the potential of the school environment to tackle issues of disaffection in a systematic way is consistently highlighted. For example, a developmental manualized intervention has been used in traumatic loss (Saltzman et al. 2001), and attention has been drawn to the role of early intervention in externalising problems (Stormont 2002). In addition the role of the school in implementing empirically supported interventions is recognised (McIntosh, Rizza, & Bliss 2000). Significant contributions reviewed in following sections of this chapter include the Conduct Problems Prevention Research Group (2002b), Weiss et al. (1999), and Reynolds & Robertson (2003).

In this discussion of the role of schools in providing a setting for interventions for a range of target behaviours for children, it is relevant to consider the effects of schooling itself on those behaviours. There is evidence for a causal relationship between school achievement and self-esteem (D'Amico & Cardaci 2003; Clark, Prior, & Kinsella 2002), and achievement has been identified as having significant role in social function (Chen, Chang, & He 2003), friendships and peer acceptance (Wentzel & Caldwell 1997). Despite the evidence that concentrating on academic enrichment and social skills does not produce the breadth of gains claimed by other approaches (see above), the effects of academic achievement on self-esteem, self-worth and self-concept are robustly

reported. Reading problems are associated with low self-esteem (Lindsay et al. 2003), and the negative effects of low performance begin early (Chapman, Tunmer, & Prochnow 2000; Bouffard et al. 2003). Low reading achievement is linked to attention problems (Rabiner, Coie, & Conduct Problems Prevention Research Group 2000), and reading ability is highly predictive of academic self-concept (Chapman & Tunmer 1997). Impacting on self-concept is seen as vital in resolving school troublemaker-victim situations (Marsh et al. 2001), while other studies have proposed a link between cognition and motivation in boosting academic performance (Kreitler et al. 1995). There is also evidence that unfocussed praise may be detrimental to motivation, while sincere praise is associated with improved performance (Henderlong & Lepper 2002). Controlled studies show a beneficial effect of counselling (Sherr & Sterne 1999), and an impact on mental health service usage for some intervention groups (Slade 2002). School-based studies have also reported on the effect an Intervention has on the school system (Gottfredson, Jones, & Gore 2002), effects on specific behaviours (Barrera et al. 2002), and using different group intervention techniques (McArdle et al. 2002). These pieces of evidence are important in consideration of the role of the school because they point to the positive elements of school practice that create an effective context for other interventions, particularly those to impact positively on child mental health.

Section 1.3: Coping with the Complexity of Disaffection

There is a demonstrable case for preventative, targeted group Interventions based in ecological settings. But what to do with the disaffected, having gathered young people together in accordance with the arguments made above? What does one have to do to be effective in reducing disaffected behaviours? The centrality and enormity of the question is made perplexing by the range of symptoms and severity that inevitably arises in an intervention targeting the disaffected (when the term disaffection is taken to denote all those at-risk of disturbance). The at-risk group will by definition include those without a clinical diagnosis and without an acknowledged treatment pathway. What does one have to do to be effective in reducing those at-risk behaviours?

To carry out outcome research the literature encourages the categorisation of conditions in a reliable and valid way, for the obvious reason that research findings need to be generalised. Research findings encourage confidence in asserting that, for example,

depressed people are most likely to benefit from cognitive therapy, or that research has shown that 90 per cent of people with panic attacks will recover with anxiety management. But how can outcome research guide what to do when with a group of disaffected children? It does not take too much imagination to see that teasing out behaviours in order to ape the discrete treatment of clinical sample and form groups of at-risk depressives and at-risk conduct problems could be tricky. With symptoms not meeting full DSM-IV criteria for diagnostic category, selecting out discrete groups based on those categories will by definition be a hit and miss process. A group of disaffected young people is inevitably and undeniably a complex mixture of symptom, physiology and psychology and this brings a complexity to dealing with these groups.

This uncomfortable fact is a bit of a black hole for outcome research into disaffection: children do not 'have' disaffection like they have the measles; what they have are behaviours and experiences. Writing about psychotherapy, a recent article picks up this theme '...if one cannot validly categorise experiences as illnesses, how can we generalise what we find in any research study? If so-called symptoms spread across diagnostic categories like spilled ink flowing over paper, how do we reliably differentiate between conditions?' (Marziller 2004 p392). The point is that where there is no clear definition of illness category (in the form of a DSM-IV diagnosis), there is no clear direction from the evidence-based tradition to guide intervention.

The vast majority of studies will test the effects of a conceptually distinct treatment modality (e.g. CBT) with a diagnostically distinct symptom cluster (e.g. depression), giving an array of sub-varieties of treatment and symptom with a competing (and often confusing) set of outcome profiles under which some sort of outcome is usually present for most treatments. The result is often that practitioners will make choices according to an idiosyncratic set of preferences and prejudices according to what they feel comfortable doing (Murray 2000). Operating in the piecemeal way that the literature expresses itself does not give proper account of the overlap between symptom cluster that characterises 'real-life' disturbance (i.e. the difficulty in finding a 'pure' group) nor does it account for the multi-problematic nature of disturbance (i.e. the eligibility of a single group member for several diagnostic groups) and such a piecemeal approach gives scant recognition of positive approaches to treatment. Even when groups more specific than those at-risk are considered... "People who come for help for a specific

problem almost invariably have other problems often to do with their personality, social conditions, emotional experiences or way of life” (Marziller 2004 p393).

This confusion can be traced back to a set of clinical guidelines called The Diagnostic and Statistical Manual of Mental Disorders and first published in the 1950s. To this day it provides an organisational structure for abnormal psychology and psychopathology. The roots of this collection can be traced to Emil Kraepelin, a psychiatrist graduating in the late 19th century and the author of the *Compendium of Psychiatry*, a book that was to have a huge impact on the theory of psychiatry. In the *Compendium*, Kraepelin set out the grounds for the central assumption that there exists a discrete and discoverable number of psychiatric disorders. Although he recognised that some symptoms occur in more than one disorder, he argued that each disorder has a typical symptom-picture. He also believed that the different disorders were associated with different types of brain pathology and with different aetiologies. On this view, the step towards discovering the causes of mental illness was to identify the different disorders on the basis of their symptoms. The diagnostic systems currently advocated by the World Health Organization and the American Psychiatric Association (the ICD-10 and the DSM-IV respectively) are similarly organised in a way that reflects Kraepelin's assumptions about the nature of madness. The science of psychiatry has been built largely on these assumptions, and while recognising the benefits brought to many people through the work of medical teams in diagnosing and treating mental health problems, it is undeniably incomplete as a system. When for example the basic test of inter-rater reliability is applied to the diagnostic system, the efficacy of the system is at best questionable (reliability concerns the consistency with which diagnoses are employed by different clinicians or on different occasions). Bentall presents a persuasive collection of evidence on this point; using the examples of schizophrenia and bipolar disorder he concludes:

‘In this chapter I have drawn on a wide range of research. None of the findings we have considered supports Kraepelin's diagnostic system. Studies of patient's symptoms, the role of genes, of the course and outcome of illnesses over time, and of the response of symptoms to treatment, all point to similarities between schizophrenia and bipolar disorder patients, rather than to differences.’ (Bentall 2003 p94)

The fundamental idea behind this analysis of the DSM-IV system is the proposition that it is more productive to try and explain and understand the actual experiences and behaviours of people with mental health problems instead of relying on psychiatric diagnosis: a so-called 'illness ideology' that defines mental health as concerned with deviant, maladaptive conditions with organic origin within the individual (Maddux 2002). Bentall writes from the position of one working with psychotic people, but the distinctions he makes are valid when relating to disaffected youth. In his view, such experiences and behaviours can be called complaints, understood as any class of behaviour or experience that is singled out as sometime troublesome and therefore worthy of attention, and symptoms understood as behaviours that are clustered together in order to make a diagnosis. The important point to be considered here in relation to the development of effective intervention programmes is that once complaints have been explained... 'there is no ghostly disease remaining that also requires an explanation. Complaints are all there is.' (Bentall p141).

The notion of studying complaints and not relying on psychiatric diagnosis is hardly new. It was anticipated by Adolf Meyer, but is thought to be first clearly articulated by the British psychologist Donald Bannister in the late 1960s (Bannister 1968). It is not new, but Bentall is chosen here to represent a swell of contemporary opinion to support approaches that build on assumptions other than those shared by the DSM, OCD and so-called medical thinking. Bentall summarises these alternative assumptions in what he calls a 'post Kraepelian manifesto' (p143). Of particular interest to the development of clear behaviour outcomes that help to form the guiding principles of the AGI intervention in youth disaffection is the emphasis in this manifesto on the social. Bentall urges us to consider complaints as endpoints of developmental pathways, which are determined in part by environmental processes. Further, that there is an aetiological role of social and biological factors in the mechanisms of psychological disturbance.

The persistence of the DSM paradigm in the face of considerable negative evidence reflects that difficulty in finding an alternative. Bentall is doubtful about the chances of doing so:

“Although we cannot entirely exclude the possibility that a future Linnaeus of psychiatry will achieve what others have so far been unable to accomplish, I am sceptical whether any new system will succeed where Kraepelin’s has failed.” (Bentall p141)

The critique of DSM thinking has benefits for the development of the intervention models used in the AGI study and outlined in Chapter 3. It allows the focus of intervention to move from diagnosis to behaviour and experience. The emphasis on understanding the actual experiences and behaviours as an alternative model has many attractions for those concerned with behaviour in school. Firstly, using a social aetiology in preference to (but not to the exclusion of) a medical aetiology empowers key social actors to frame change. Secondly, the notion that the diagnostic entities for treatment are not reliable means that behaviours need not be framed exclusively by diagnostic category, or by a deficit model of dysfunction. In other words it opens the door for the development of a positively framed structure for desired behaviours in youth disaffection. As this challenge is taken up in Chapter 3, it will be useful to remember that the position of those responsible for the Action Groupskills Intervention are guided by the insight that over-reliance on diagnosis can act as a short-cut to a simplistic application of treatment procedures where the complexity of behaviours demand a fuller account of experience and the skills-gaps that are being manifest through behaviours.

When the outcome literature is systematically reviewed (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002; Rutter, Giller, & Hagell 1998), helpful trends in the development of therapeutic approaches are observed - helpful, that is, to approaches seeking to compliment DSM thinking. Fonagy isolates trends that are an implicit recognition both of the proliferation of forms in evidence-based practice, and the inter-relatedness, that is to say complexity, of disaffected behaviours. The trends are: the developmental framework and the merging of preventative and treatment approaches, the moderation of the radical goals of treatment, the decline of generic therapies and emergence of specialist treatments, multi-component interventions and multidisciplinary work, sensitivity to contextual effects and individual differences, and finally, user empowerment. These are trends in outcome-based practice woven through the literature in child and adolescent treatment, and are evidence first of “...the greater sensitivity to developmental considerations in the design and implementation of both physical and psychosocial treatment strategies, both in terms of coupling physical and psychosocial treatments and

in linking different treatment modalities (behavioural, cognitive, systemic etc.).” (Fonagy p393) Secondly there is evidence of increasing awareness of having to consider the system within which the child functions (the family, the school, the community, etc.), leading to a “heightened interest in developing packages of interventions for different people in the dysfunctional system, using different treatment approaches and involving multiple agencies as service providers. It has also produced increasing concern about the child’s social interactions and social functioning rather than simply symptomology.” (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002 p394).

The picture that emerges is remarkable because it emerges from the much maligned ‘medical model’: It paints childhood disaffection as having many forms and multiple causes all needing an account in a coherent and consistent theoretical model that links different models of ‘pathology’. Behaviours must be understood in the context of the environment and changes in behaviours should be mirrored by changes in the environment. In this respect the analysis offered comes close to conventional critiques of the syndrome model of childhood disturbance that emphasise the constructed nature of behaviours (see e.g. (Laslett 1983). Academics and practitioners have not been slow to herald a shift in locus of expertise from physiological psychopathology to developmental psychopathology, from medical psychology to sociology (Jones 2003). The message to preventative practitioners is that treating discrete diagnostic categories is a complex enterprise, and that the extended attention received by the question of what works for whom has revealed trends in the way this complexity is managed. The trends in treatment tend to favour heterogeneous approaches based on integrated or pluralistic understandings of childhood dysfunction proposing systemic change to meet practical ends. This is good news for practitioners who are led to complex treatments approaches by the complexity of the presenting behaviours. The intuitive rightness of pluralistic approaches is in rare agreement with evidence based literature and it is relatively simple to extrapolate from these trends to guide effective preventative practice. Overarching and pluralistic theoretical frameworks, and as a consequence multi-method systemic approaches in assessment and treatment will characterise interventions accountable to the challenge of complexity.

The model that emerges from this argument draws inspiration from a combination of biological, systems, cognitive behavioural and psychodynamic perspectives. Evidence

for studies employing approaches informed by dual frameworks (or tri, or even multi-framework) is more easily inferred from attitudes than explicitly stated, but it has been consistently noted in the literature e.g. (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002) who believes "this approach is a response to the burgeoning research literature on the naturalistic and experimental treatment outcome of childhood mental disorder", going on to say that "...a quarter of a century's research in developmental psychopathology confirms the view that specific problems are likely to be the consequence of heterogeneous causal determinants including the transactional interaction of biological predisposition with lived experience." (p394). Psychosocial interventions in particular, but perhaps also medical interventions are changing their foci in line with an implicit systemic model where theoretical frameworks are subservient to the higher aim of fully accounting for the behaviour of the individual within the environment. What is exciting in itself, but of particular importance to this study of the effects of therapeutic groupwork in schoolchildren, is that adhering to the precepts of a particular framework (e.g. psychodynamics) is less important than effecting lasting change in the individual, and to the end of effecting change, frameworks can be modified and enlarged. The change is perhaps clearest in the evolution of behavioural treatments. These have moved from an approach firmly rooted in positivistic learning theory that denied the importance of all processes beyond those entailed in classic and operant conditioning, to an orientation that, at least implicitly "...recognises the importance of the child's and parent's feelings and thoughts (emotions and cognitions) as determinants of behaviour. This shift has quickly led to a broadening of CBT interventions to include disorders that are principally affective in nature (e.g. depression). More important in our view, among clinicians within this orientation there is increased concern with the emotional environment of the child." (Fonagy, 2002 p396), and sometimes does so explicitly; (Howard & Kendall 1996; Meichenbaum 1997).

Consideration of the experience and genesis of emotions, as distinct from the role they play in social conditioning goes against the founding principles of behaviourism and includes communication patterns in the family (Gottman, Katz, & Hooven 1997), which have thus far been of predominant concern of family therapists. The example set by the theoretical mobility of behaviourism is echoed increasingly in other theoretical approaches. In his consideration of the future of adolescent psychotherapy groups,

MacLennan reflects on the increased development of short-term problem-focused groups and the use of cognitive behavioural approaches to treatment:

“[this] has been accompanied by the increase in structured groups, which include behavioural reward systems, role-playing and journaling, the use of well-defined contracts and the development of a broad range of structured materials, games and manuals related to specific problems”. (MacLennan, 2000 p70)

The author goes on to express reservations about the efficacy of short-term groups using structured cognitive behavioural methods and focused on a limited range of skills or problems, but the principle of incorporating theoretical approaches that might be considered inimical into programmes evaluated for comparative effectiveness has already been accepted.

This characterisation of the literature presents trends in practice that use the outcome research to guide interventions to meet general problems to do with personality, social conditions, emotional experience and way of life in the target group. These interventions are guided by the nature, severity and individual character of the problems encountered and may well borrow techniques across theoretical frameworks. The ‘borrowing’ that takes place applies to most target groups, but for reasons discussed above is particularly relevant to at-risk populations.

Practitioners and theorists have looked at these trends to reflect on the priorities and guiding principles of the content of intervention (what to do when?). In doing so there is an acknowledged need to bring order to those areas where treatment frameworks compete rather than complement each other. There is a feeling that too often school interventions are missing the components necessary for making accurate interpretations of the intervention outcomes (Lane, Umbreit, & Beebe-Frankenberger 1999), and that to be more successful, the content might be organised around overarching themes. There is surprising commonality in recognition of the issue (what guides intervention content?) and the themes used to guide intervention. Recent studies that frame their investigation in terms of developmental domains such as emotional competence (Denham et al. 2003; Lengua 2003; Orobio de Castro et al. 2003; Pierrehumbert et al. 2002) play (Leblanc & Ritchie 2001; Hansen, Meissler, & Ovens 2000), social ability (Adalbjarnardottir 1993;

Camodeca, Goossens, Schuengel, & Terwogt 2003), and attachment (Diamond et al. 2002; Stilwell, Glavin, Kopta, Padgett, & Holt 1997), are orientated around using intervention approaches directed towards effectiveness in delivering change.

The characteristics of these effective interventions are summarised by (Rotheram-Borus & Duan 2003) as investigator-driven, theory-based, focussing on changing the target behaviour and implemented with fidelity over time. These characteristics are echoed in another highly relevant paper (Lane et al. 2001), that makes the case for social validity, treatment integrity and generalisation and maintenance to be addressed in designing effective interventions. These 'meta themes' guiding content to effectiveness build an approach that is rooted in theory, has one eye to the outcome research in as far as it is useful in guiding the content of interventions, but is orientated towards the individual in the group and providing real-life outcomes. These themes emerge again as the need to demonstrate outcome is examined.

Section 1.4: Coping with the Need to Demonstrate Outcome

The issue of whether particular treatments for disaffected and disturbed children actually works is not a new one for practitioners seeking to alleviate the distress caused by the symptoms of disturbance. What has changed over the years is the nature of the evidence deemed acceptable in demonstrating the effectiveness of treatment approaches. The current zeitgeist in this area is can be characterised by the rise in 'evidence-based practice' which is "... at its core all approach to knowledge and a strategy for improving performance outcomes" (Alvarez & Ollendick 2003). Evidence based practice is not wedded to any one theoretical persuasion or orientation, but holds that treatments of all characters need to be based on objective and scientifically credible evidence.

The credibility of some treatments has received some considerable damage in the past, as the evidence for effectiveness has been assessed. Famously (Eysenck 1952) and (Levitt 1957) concluded that psychotherapy and child psychotherapy were no more effective than the simple passage of time. In response the professional community developed methodologies more amenable to outcome assessment resulting in over 1500 studies (Durlak et al. 1995; Kazdin 2000) and major meta-analyses examining the

effects of child psychotherapy (Casey & Berman 1985; Kazdin et al. 1990; Weisz et al. 1987; Weisz et al. 1995). The example of psychotherapy in expanding and developing a credible evidence base is echoed by the majority of theoretical perspectives and has resulted in the current movement in evidence-based practice. One of the most recent reviews in this area define empirically supported treatments as those being “shown to be superior to a psychological placebo, pill or another treatmentin addition effects supporting a well-established treatment should be demonstrated by at least two different investigatory teams” (Ollendick & King 2004 p5). The same review holds that scientifically credible evidence will be evidence ‘obtained from randomized clinical trials (RCTs)’. In a RCT, children with a specific presenting problem are randomly assigned to one treatment or another, or to some control condition, such as a waiting list, or attention-placebo condition. This review expresses the received wisdom about controlled trials that ‘although such a design is not failsafe, it appears to be the best strategy for ruling out biases and expectations (on the part of both the child and the therapist), that can result in misleading research findings’ (Ollendick & King 2004 p4). In this matter there is agreement with other major reviews seeking to define effective practice. For example, Fonagy sets out his list of inclusion criteria for studies

“(1) group design involving random assignment, (2) well-documented treatment procedures, (3) uniform therapist training, or clear manualisation of the protocol for administering medication and evidence of adherence, (4) clinically referred samples of treatment candidates, (5) outcome assessment, including at least two outcome levels (e.g., level of symptoms, adaptation, mechanisms, transactions or service utilisation), (6) tests of clinical significance, (7) assessment of long term outcome (follow-up of greater than 6 months)” (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002 p38).

There are however some reported difficulties in using the RCT as the central arbiter of evidence-based practice; namely that “only 7.4% of studies in child and adolescent mental health met [the criterion of] rigorous randomisation” (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002 p38). If the evidence-based movement is to be based on 7.4% of studies then it could more accurately be described as a cult than a movement, the real movement being in the character of the remaining 82.6% of studies. A pragmatic approach to this practical difficulty in establishing a revolution with patchy numbers for inclusion is to adopt a hierarchy of evidence (e.g. Clarke & Oxman 1999), that allows

rigour to be preserved at the same time as admitting more evidence for analysis. There are, however, other philosophical problems with the RCT, as brought out in the debate about the desirability and utility of RCTs for obtaining “reasonable evidence” (Persons & Silberschatz 1998). And while it is still possible in the face of such reservations about the RCT to make a strong case for accepting that such trials are well suited for establishing *initial* efficacy (Ollendick & King 2004), there is an imperative need to demonstrate the transportability of the treatment to practice settings and their effectiveness in the ‘real-world’.

It will be helpful to distinguish at this point between the methods considered acceptable for generating outcome and the question of what is ‘outcome’? As far as the question about method is concerned it is right to be wary of the pre-occupation with randomised controlled trials that exists in the literature. At the same time it is right to be cautious of those who disown generalisability but are less clear about what stands in its place (Salmon 2003). The *Cochrane Reviewers Handbook* (Clarke & Oxman 1999) sets out broad categories of method in a hierarchy to distinguish studies according to their susceptibility to bias: (i) randomised controlled trials; (ib) systematic reviews and meta analyses; (ii) other trials: a controlled trial without randomisation, a quasi experiment, or a failed randomisation; (iib) experimental single case designs; (iii) cohort studies preferably from more than one centre (a cohort allocates by exposure to treatments and looks for differences in outcomes); (iv) case control (retrospective) studies (allocates by outcome and looks for difference of exposure – in terms of treatment; (v) large differences reported in comparisons between times and/or places, with or without interventions; (vi) opinions of respected authorities based on clinical experience, descriptive studies, uncontrolled studies, and reports of expert committees.

In order to have a claim to demonstrate outcome, studies must have a sensitive design for evaluation that is persuasive to the target audiences and must therefore take steps to eliminate bias from their reports while being realistic and open about the bias still remaining. To define good outcome as performance in a controlled trial is not the intention of evidence-based practice, but in order for outcome to be considered seriously researchers are expected to follow rules that make their task difficult and distinguish social science from journalism. It is inescapable that research is judged against the

conventions of discipline even when acknowledged that the level of discipline is judged for its own sake (Salmon 2003).

Aside from the questions of method in the need to demonstrate outcome there is a need to be clear about what is “outcome”? Or whose outcome is it anyway? Fonagy draws our attention to the tendency to regard child mental health outcomes in absolute terms, when fundamental questions with practical implications are rarely asked: “...when is the preservation of the family unit a positive outcome, or, rather, for whom is that outcome positive – the child, the parents, the clinician, or the purchaser who would be required to fund alternative care? What if the outcomes diverge – if what is optimal for the child is less favourable to other members of the family or to the service provider? We should be aware of the adage: ‘In many instances the most cost-effective intervention is to do nothing.’”(Fonagy, Target, Cottrell, Phillips, & Kurtz 2002 p4). The question of ‘for whom the outcome is a positive one?’ is exacerbated by the divergent view of the participants of the nature and severity of the condition; even at the level of symptoms, teachers, mothers, and fathers appear to share little agreement (10% of variance) concerning their perceptions of Internalising symptoms (Achenbach 1995). With this in mind, establishing agreement concerning good outcome looks challenging. Even with the outcome between participants agreed can there be a valid concept of good outcome across cultures and ethnic groups?

Psychotherapy apparently short-circuits the culturally constructed nature of ‘good outcome’ through the emphasis on the achievement of selfhood through the separation-individuation process (Mahler 1971) as one of the cornerstones of psychotherapeutic intervention. And yet is it not correct that the emphasis on individual achievements in Western culture is excessive, and that the appropriate submission to the goals of the family and community may be a better indicator of healthy adaptation (Lasch 1978)? Interventions ought to be consonant with the subjective culture of the ethnic group to which they are applied, and the instruments used should be able to integrate cultural meanings with the pertinent scientific categories, but are these concerns consistently referenced in the growing literature of the evidence-based movement?

Against these concerns it is easy to hear the voice raised with impatience: ‘but surely common sense will prevail and good outcome will be demonstrated by a reduction in

symptoms?' Objective suffering should be able to give us the immovable point from which to judge good outcome. But even here it is not so simple: even overlooking the fact that objective suffering is inferred from reported data and questionnaires i.e. is unavoidably secondary, the developmental framework that has emerged to dominate child psychiatry and psychology (Cicchetti & Toth 1995), implies that the symptoms cannot be considered the sole, or even the most important, criterion of treatment effectiveness. Fonagy puts it well when he says: 'if psychiatric disorder is not just the end result of a series of interactions of biological, social and psychological characteristics across time, but is itself part of a complex transactional causal chain, good outcome might sometimes be an increase rather than a decrease in symptomology' (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002 p5). These points are raised not to denigrate the need for outcome, nor to cast a slight on the evidence-based movement, but to demonstrate that evidence is not absolute. Inevitably its significance is determined by the cultural context that demanded it and gives it meaning. With this in mind it becomes essential to consider evidence from within this relativist framework.

Section 1.5: Analysing the Evidence of Research Programmes.

The discussion in sections 1.1, 1.2, 1.3 and 1.4 has set out the theoretical case for components to be included in an effective intervention with disaffected youth. The components can be summarised as:

Issues of Scale

- A targeted prevention focus
- A group intervention approach
- A non-clinic, real-world setting

Issues of Complexity

- A systemic understanding of disaffection (the contributing role of the environment)
- Intervention method based on an integrative analysis of disaffection
- Content that reflects 'meta-themes' of effectiveness

Issues of Outcome

- **Rigour in methodology and commitment to outcome**
- **Outcome measures from multiple viewpoints**
- **Awareness of cultural determination in issues of outcome**

These components are considered to be important in delivering an effective intervention, and in this section they are used to structure an examination of the outcome literature in the area of disturbance. Studies relevant to the issues above are examined in an order that is roughly based on the order of components above, but because studies have been sought for their ability to combine a number of components the discussion quickly moves to detailed examination of central studies that cross over between components. The purpose of this exercise is to demonstrate the good sense of designing an intervention like the AGI study by referring to similar studies containing some of the same components that have provided positive outcomes.

- **The first sub-section 1.5.1 deals with the prevention component: this being the first distinctive component identified as important in effective interventions. The evidence supporting targeted prevention is placed in the context of published prevention studies with a wider focus. Key studies are highlighted that look for a good example of this component together with other components.**
- **The second sub-section 1.5.2 deals with the non-clinic setting and group component: looking particularly at interventions active in the school setting. Key studies are highlighted that are good examples of this component together with other components.**
- **The third sub-section 1.5.3 deals with the complexity component: comparing outcomes from contrasting theoretical frameworks. Key studies that provide comparative data are highlighted as good examples of this component together with other components.**
- **The fourth sub-section 1.5.4 begins with a table to illustrate how the highlighted studies satisfy the desired components. A more detailed examination of these key studies is offered.**

1.5.1: the prevention component

As reported above, targeted prevention approaches are rare in outcome literature, mainly because of the problems in demonstrating outcome in sub-clinical samples. Because targeted, or secondary, prevention is framed by the wider notion of primary prevention, it is reasonable to consider the effectiveness of these wider programmes in order to support the case for the targeted prevention approaches, where the literature is a little thinner. Primary prevention has a long and distinguished history in America through the Headstart Program. McKey et al presents a meta-analysis evaluation of the programme initiated in the 1960s with young children living in disadvantaged circumstances in which immediate effect sizes relating to cognitive aptitude and social behaviour are in the order of half a standard deviation (McKey et al. 1985). Although these are seen to tail off at the 3 year + follow-up point, effects are still present. Lazar and colleagues were able to demonstrate the effectiveness of the more intensive and well-documented controlled interventions in the Headstart programme by way of synthesis of 12 studies (Lazar & Darlington 1982).

The findings strongly suggest that, compared to controls, those children who attended programmes were less likely to be referred to special classes or to repeat grades and although the IQ gains tended not to persist, there were longer-term benefits in social functioning. The High/Scope Perry pre-school study stands out with respect to its focus on high risk families, the quality of its programme, a low attrition rate and the length of its follow-up to 27 (Schweinhart 1999; Schweinhart & Barnes 1993; Weikart & Schweinhart 1992) The longitudinal sample consists of 123 aged 3 and 4 years old identified by survey and included because their families lived in poverty (as assessed by parents' years of schooling and employment status) and because the children scored low on intelligence tests given at age 3. The internal validity of the study rests on the random assignment to groups: a pre-school group receiving developmentally appropriate curriculum devised and delivered by teachers and a no pre-school group. The pre-school group met weekday mornings for 30 weeks, during which time the teachers made a weekly visit to the home, working with parents for 90 minutes on supporting and modelling the intervention made in the pre-school environment. The children were assessed through parent interview, annual intelligence tests, annual achievement tests, annual teacher ratings, participant interviews and police & social services records information collected at age 19. The findings reported are those whose probability of

chance occurrence is less than $p=0.05$, and the pre-school group consistently outperformed the no-pre-school group on intelligence tests up to the age of 7, but showed no difference from then on. Achievement T-scores showed no difference between the groups until scores taken at 14 and 19, when the pre-school group outperformed in literacy and subscales relating to health knowledge. Employment, involvement in misconduct and police arrests distinguished the groups in favour of the pre-school group. In an exercise to demonstrate the worth of this programme in economic terms, it was calculated that for every dollar invested, the taxpayer received a return of \$5.95, in reduced costs for special education, crime costs and welfare payments.

The Perry Preschool project provides persuasive evidence of the effectiveness of a substantial intervention in the pre-school years. The findings are hard evidence in that they do not allow conclusions to be drawn about the nature of disturbance and the natural history of its development, but rather point towards outcomes in activities likely to be undertaken. The small size of the sample and the conservative nature of the statistical test allows the finding of significance to point to an effective intervention. The intervention itself is of significant length and complexity to be costly, a criticism anticipated by the cost-benefit calculation, a luxury afforded by the cohesiveness of the experimental design and the low attrition rate. These findings are broadly echoed in a meta-analytic review of primary prevention mental health programmes for children and adolescents conducted by Durlak: (Durlak & Wells 1997). 177 programmes are reviewed using meta-analysis and found to provide empirical support for practice. In practical terms the average participant in a primary prevention programme surpasses the performance of between 59% to 82% of those in a control group, and outcomes reflect an 8% to 46% difference in success rates favouring prevention groups. In an earlier review Yoshikawa found successful primary delinquency prevention achieved through early family support and education appears to have a number of common factors (Yoshikawa 1994). Programmes that achieved long-term reductions in antisocial behaviour and delinquency also addressed multiple risk factors in families where children were between the ages 0 – 5. The Perry Pre-School Project, The Syracuse Family Development Research Project (Lally, Mangione, Honig, & Wittner 1988) and the Yale Child Welfare Project (Seitz, Rosenbaum, & Apfel 1985) all had effects on multiple risks, an ecological multiple setting design and a length of at least 2 years.

Primary prevention programmes can be effective but also by necessity are expensive and broad in application. They must include children not at-risk and the interventions themselves must be benign enough to be acceptable to all children and their caregivers. The sensitivity of the Headstart programme to criticism on the basis of economics, evidenced by the analysis of cost-benefit, is an example of how prevention programmes like this must justify their complexity and expense (...and like all complex programmes justifying themselves through simplistic cost-benefit analysis, they warrant a little scepticism). The categorical data of these programmes is non-parametric and although the hard evidence of the outcomes is attractive, and indeed a strength of the programme, the more subtle analyses about the nature and development of anti-social behaviour are lost. The statistical significance is derived from comparison of intervention and control groups (i.e. the intervention is more differentially effective than the control) but ANOVA analysis to see where the difference occurs across time and condition is not applicable because of the nature of the data. The main problem with this is the specific effect of the intervention is not identifiable: it is only inferred through reference to categorical data (who is in work, who is out of care etc). Therefore replication in other situations and cultures becomes difficult because specific problems are difficult to anticipate and because of the expense of the project, the risks become prohibitive.

Key Studies:

Focussing the intervention on children screened and identified as at-risk resolves some of these difficulties. Intervening at the early onset of anti-social behaviour is categorised according to the criteria above as secondary prevention, and typically operates as follows. A particular population is screened or evaluated in some way and criteria are used to target some for intervention, which follows quickly after the collection of more information confirming the nature of the children's difficulties. The intent of secondary prevention is to help children with sub-clinical problems so that they avoid developing full-blown disorders. It is believed that the earlier the intervention occurs, the greater the likelihood of success. In other words, it makes sense to intervene when problems are just beginning rather than wait for them to intensify over time. One of the largest in scale is the Primary Mental Health Project (PMHP) begun by Cowen in 1957 (Cowen, Hightower, Pedro-Carroll, Work, Wyman, & Haffey 1996). So called because of its focus on the primary school years and not because it is a primary prevention project, the

PMHP offers non-behavioural treatment to children who have Internalising or Externalising problems as well as learning difficulties. A more thorough and convincing study has been published by Reynolds and distinguished by methodology and relevant outcome data (Reynolds & Robertson 2003). This study receives detailed consideration below.

1.5.2: the non-clinic setting and group intervention component

The importance of non-clinic settings for interventions has been made above and can be summarised as follows:

- i) Many children and their families who suffer psychological distress are reluctant to attend a clinic. This will be particularly so in those cases where behaviours are a cause for concern but considered to be at a sub-clinical level. The reluctance to submit to the stigmatisation of clinical referral will impede preventative interventions in a clinical setting.
- ii) The problem may be relatively straightforward and the mobilisation of the full range of expertise of the clinic base not needed.
- iii) In recent years there has been enormous change in the policies and practices of social agencies for children, principally education and social services agencies in response to new demands such as the extent and effects of child abuse alongside changes in social pressures and political outlook.
- iv) Children's behaviour is very much determined by the environment they are in at the time. That behaviour and emotion arise within a social context is demonstrated by numerous studies e.g. (Jones 2003). If feelings are detached from the ecological setting, much of the potential for effective intervention may be thrown away.
- v) In addition, the results of treatment may be situation specific, examples of successful treatment that have not generalised to environments other than that in which the treatment took place are not rare. A partial answer is to take the treatment to the environment, be it home or school where the problem is manifest.

These statements of good sense are supported by lack of comprehensive evidence from clinic treatment outcomes (Sourander & Piha 1998), (Grizenko 1997) and solid long-term data from parenting programmes in the community (Lally, Mangione, Honig, & Wittner

1988), and more focussed work in the clinic setting but of a non-traditional nature (Patterson, Reid, & Dishion 1992; Webster-Stratton, Reid, & Hammond 2001). Whereas the non-ecological setting of the clinic may add validity and reliability to short-term treatment gains, the limits it imposes in terms of long-term gains and generalisation of gains subject it to some of the same ambiguous findings of more traditional clinic-based treatments. Interventions in the school setting enjoy the uniqueness of setting and socialising environment that translates well into cross-contextual gains over time (Lochman 1992; Kolvin, Garside, Nicol, Macmillan, Wolstenholme, & Leitch 1981). Other effective interventions in non-clinical settings include that of Kataoka, Stein, Jaycox, Wong, Escudero, Tu, Zaragoza, & Fink (2003) and it is noted that there are fewer preventative interventions located in non-clinic settings than might be expected considering the numbers of participants involved and the logistical difficulties in ensuring attendance at clinic-based sessions (Reynolds & Robertson 2003).

The importance of the school environment has been highlighted in section 1.2. As mentioned there, schools-based outcome research is a little patchy, but good studies do exist. For example, Durlak drew conclusions about the efficacy of school-based prevention programmes (Durlak 1995). Relying upon meta-analysis of 130 controlled outcome studies published and unpublished, appearing before 1991 (Durlak & Wells 1994), he reported that programmes targeting Externalising problems achieved the highest effects (mean ES = 0.72), while the most common Internalising problems of anxiety and depression, appear amenable to early intervention (mean ES = 0.49). Although the effects for children with academic problems and poor peer relations are more modest (ESs of 0.26 and 0.30 respectively), each of these dimensions is predictive of later adjustment so that even modest changes on these indices can have some preventative effect.

Kolvin screened 4300 children in the North of England to select 600 'at-risk' children showing signs of having social or psychiatric disturbance or learning problems (Kolvin, Garside, Nicol, Macmillan, Wolstenholme, & Leitch 1981). Predicting that there would be differences in the impact on at-risk behaviours between four comparable groups of children who had received different types of treatment or no treatment at all, the team found statistical differences between no-intervention and intervention groups, and differences between the types of intervention on offer (see below). All of these

interventions were school-based and there was evidence of cross-contextual gains, in addition, treatment effectiveness showed an increase with time (p305). Lochman and his colleagues reported that a cognitive-behavioural programme was more effective amongst boys identified as aggressive by schoolteachers than either goal-setting alone or no treatment in reducing disruptive aggressive off-task behaviour in the classroom (Curry, Wells, Lochman, Craighead, & Nagy 2003; Dunn, Lochman, & Colder 1997). The addition of goal-setting to the cognitive-behavioural intervention resulted in greater reduction than the cognitive intervention alone (Lochman 1992). Another recent study found significant treatment effects at a 24 month follow-up for groups of children assigned to a school-based treatment group or a monitoring group (Dadds et al. 1999). This study evaluated the long-term effects of a combined child and parent focussed intervention for prevention and early intervention for anxiety problems. A devised scale of diagnostic ratings alongside a ratings of change scale were administered by clinicians unaware of treatment status in a telephone interview with parents at 6 months, 12 months and 24 months. The resulting diagnostic severity ratings of the clinician were used to determine percentages of children in each treatment group meeting criteria for anxiety disorder (DSM-IV) and the parent rating of change used with a separate clinician rating of change to generate comparison data suitable for a multivariate analysis of variance. The MANOVA evidenced a treatment effect at the 24-month stage that was not present at the 12-month follow-up. The categorical data found differences between groups in rates of diagnosis emerge at 6-month, converge at 12 months and re-appear at 24 months. The obvious methodological problems with the study are the validity and reliability of the devised scales used, the fact that a clinician is making a judgement in rating children based on a telephone conversation with a parent who cannot be treatment status blind, and conclusions about the effectiveness of the intervention are made without supporting evidence from either parental questionnaire or self-report, or teacher report. Other major studies include the Montreal Longitudinal-Experimental Study (Tremblay & Craig 1990) focussing on high-risk children over a period of two years, and the Fast Track intervention model developed by the Conduct Problems Prevention Research Group (Conduct Problems Prevention Research Group 2002a). The Fast Track program is a largely successful attempt to apply a developmental analysis of anti-social behaviour to at-risk children in a combined parenting and school intervention. This multi-component, multi-site program is designed to identify high-risk children prior to entry into Year 1 of their schooling (4 years old) and address the major

deficits that lead to school failure by intervening with the child, the family and in the classroom. In this way positive outcomes from parenting and school programmes may be combined.

1.5.3: The Complexity Component

Finding programmes of intervention that combine the three core criteria of prevention, real-world and group is a challenge, but finding programmes that combine the core criteria with those of systemic, pan-theoretical, developmental and multiple domain is extremely difficult. There are good studies preventatively orientated, using group approaches in real-world settings (Saltzman, Steinberg, Layne, Aisenberg, & Pynoos 2001), but it is unclear whether the content of the intervention corresponds to the criteria thrown up by the consideration of complexity above. This could show a gap in the literature that this study seeks in a small way to fill, but it is more likely that the trend in the literature of outcome studies is just that, a trend, and examples are difficult to find because this characterisation of intervention theory is only just emerging. That is to say, the content of effective interventions may well adhere to a theoretical formulation like that proposed above, but because of the novelty, is not expressed in that way.

Rotheram-Borus found that although supposedly all effective interventions are theory-based, often the theories are not as comprehensive as the programmes, and are likely to be silent as to the processes by which change occurs. "Many efficacious interventions are more comprehensive than their underlying theories would lead us to expect. To maximise the probability of successful intervention results, many efficacious program activities are usually broader than the theory dictates" (Rotheram-Borus & Duan 2003). Three examples are given by Rotheram-Borus where the treatment outcomes outstrip those predicted by the treatment theory framework, indeed could not be expected if the treatment framework is adhered to strictly. This supports the analysis above where behavioural approaches were found to be 'borrowing' from other frameworks in order to enhance intervention results and is used here to point to the potential prevalence of programme content delivered in agreement with the criteria of complexity, despite there being little record of the theoretical base reported. This discrepancy between written account of theory and intervention practice is a grave omission, "...dissemination of efficacious interventions into real world settings may be hampered by inclusion of many activities, techniques, and strategies that go far beyond their underlying theories.

Greater correspondence between intervention theory and program activities appears to be warranted to mitigate this barrier” (Rotheram-Borus & Duan 2003 p518).

As the literature is reviewed from the point of view of treatment framework, it is worth bearing the points raised above in mind. There is no shortage of evidence for most treatment approaches, and recent reviews have looked at the question of what works for whom in some detail. The difficulty is in quantifying the cross-over between the stated treatment approach and supplementary techniques from other disciplines used in practice. Furthermore, there is the task of teasing out the effective elements from the resulting practice (Lane, Beebe-Frankenberger, Lambros, & Pierson 2001). What follows is a close look at the evidence from a few studies that use a particular framework as their starting point, followed by an analysis of how converging techniques from different treatment frameworks cross over to confound treatment comparison.

For example, the Fonagy review (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002) found evidence to support cognitive-behavioural interventions among all groups, behavioural interventions with types of anxiety disorder, and family interventions with disruptive children, even where cognitive-behavioural interventions have failed. This review presents little evidence for the effectiveness of psychotherapies, except with more generalised anxiety (p300ff), but reserves favourable comment for multi-modal school-based interventions across types of disorders (p315-319). These findings could be more representative of the biases in publication than the comparative effect of intervention approach. However, a recent study using a randomised control design and methodological rigour confirmed the ineffectiveness of traditional child psychotherapy (Weiss, Catron, Harris, & Phung 1999).

Ollendick and King (2004) made an evaluation of their previous review of empirically supported treatments for children with phobic and anxiety disorders and arrived at more specific conclusions about efficacy (Ollendick & King 1998). This review addresses the following procedures: systematic desensitisation, emotive imagery, modelling, reinforced practice, verbal self-instruction, cognitive-behavioural interventions and integrated cognitive-behavioural plus family-based procedures. This review asserts that “...firm empirical support for other approaches is lacking” (2004 p9), a conclusion consistent with meta-analyses of 108 treatments between 1970 and 1985 (Weisz, Weiss, Alicke, &

Klotz 1987). A more recent meta-analytic review of 150 studies published between 1967 and 1993, (Weiss & Weisz 1995; Weisz, Weiss, Han, Granger, & Morton 1995) concluded that behavioural treatments were more effective than non-behavioural treatments regardless of child age, therapist experience, or treated problem. However, the Ollendick review finds only participant modelling and reinforced practice to enjoy well-established status for treating childhood fears and phobias, whereas for anxiety disorders no treatment is well-established (Ollendick & King 2004 p9). This is hardly encouraging if other treatments are ineffective as is claimed. The limited scope of this review might well be questioned as well as the criteria upon which the studies were compared. It is not clear whether sufficient care has been taken in this review to standardise outcome data, and no mention is made of effect size. Instead the analyses appear to depend on categorical comparisons.

Despite the limitations of these reviews it does serve to illustrate the difficulty in establishing the effectiveness of any particular approach, something echoed in a controlled trial of a brief cognitive-behavioural intervention with depressed adolescents (Wood, Harrington, & Moore 1996). This study randomised 53 child and adolescent psychiatric patients with depressive disorders to a brief cognitive behaviour therapy (CBT) or a control treatment consisting of relaxation training. The treatment phase of the trial consisted of 5 –8 sessions, and post-treatment assessments showed a clear advantage of CBT over relaxation on measures of both depression and overall outcome. However there were no significant differences between the treatments on co-morbid anxiety and conduct symptoms. At follow-up, the differences between the groups were reduced, partly due to a high relapse rate in the CBT group and partly because the relaxation group continued to recover. This study, although small in scale has methodological rigour and a good design.

The findings point to the potential ambiguity behind the common assertion of the superiority of cognitive-behavioural treatment approaches over all others. Cognitive methods will be effective with certain groups at certain stages of their symptomology in certain settings. For example Lochman conducted a cognitive-behavioural intervention with aggressive boys in a school setting that had positive effects persisting at a 3-year follow-up (Lochman 1992). The design for this study used two no-intervention control groups; an untreated aggressive group and a non-aggressive group to compare with an

Intervention group consisting of aggressive boys in an anger coping programme. The programme ran weekly for 12-18 weekly sessions lasting an hour and involved children in the formation of a group culture with agreed norms of behaviour along with content focussing on the development of control mechanisms for anger triggers explored in a variety of simulated situations. Children were assessed using a mixture of interview, standardised self-report and observation procedures at a follow-up point of three years. The intervention effects were tested using multivariate analysis of covariance with the five standardised summary scores from self-report and the independent variable as condition. A significant MANCOVA condition effect was found, $F(10,230)=3.35$, $p<0.001$. Further univariate ANOVAs on the five summary scores indicated that the untreated aggressive group had significantly more negative outcomes than the non-aggressive boys, whereas the Anger Control treatment group was not significantly different from the non-aggressive group.

This study indicates that a cognitive behavioural intervention with aggressive boys can produce long-lasting effects in certain areas of their functioning. The persuasive grounds for accepting the effectiveness of this programme will include the socialising effect of the school environment in which the intervention took place, something not emphasised in the paper. Also there is an important socialising effect present in the creation of a cohesive group, regardless of treatment approach. Support for the group approach is found in a study evaluating cognitive-behavioural treatments for childhood anxiety disorders: (Barrett, Healy-Farrell, & March 2004). Children fulfilling diagnostic criteria for separation anxiety, overanxious disorder or social phobia were randomly allocated to 3 treatment conditions: 2 group cognitive-behavioural conditions, and a waiting list condition. The children were assessed at post-treatment and at a 12-month follow-up as to whether they fulfilled the same diagnostic criteria. Results indicated that across treatment conditions, 64.8% of children no longer fulfilled diagnostic criteria for an anxiety disorder compared to 25.2% of children on the waiting list.

Randomised control studies including cognitive behavioural approaches with psychodynamic-Interpersonal therapy are rare. Barkham and colleagues compared the efficacy of these approaches in a very brief intervention design (2-plus-1 sessions), (Barkham et al. 1999). Although this study was conducted with adults and uses a clinical cut off to obtain a percentage used as an outcome measure, it is worth mentioning, both

for the comparison it gives between therapeutic approaches, and the contribution it makes to the literature on brief interventions. According to stratification for severity made on the Beck Depression Inventory (Beck et al. 1961), improvement rates at the end of treatment were 67% (stressed), 72% (sub-clinical), and 65% (low-level clinically depressed) with no significant differences between treatment approach at post-treatment. A significant advantage was shown for cognitive behavioural intervention at one-year follow-up.

Of similar interest is the study published in 1995 that examined the hypothesis that the apparent superiority of behavioural interventions among children is due to differences in the methodological quality of behavioural and non-behavioural treatments (Weiss & Weisz 1995). This study used 105 studies taken from a sample identified in 1987 (Weisz, Weiss, Alicke, & Klotz 1987) using meta-analytic methodology to test whether behavioural and non-behavioural interventions differ in methodological quality and further, whether such differences account for effect size differences between behavioural and non-behavioural studies. The results suggest that using the present definitions of methodological quality, the apparent superiority of behavioural treatments is not an artefact of methodological quality.

It is noted that non-behavioural treatments have not had such extensive evaluation as behavioural and that both treatments need evaluation under conditions that prevail in clinical practice. Also, because meta-analysis is a correlational technique, the results should be regarded as suggestive rather than definitive. There are suggestions that the apparent superiority of certain treatments may be due to investigator allegiance effects. However, the results presented with thorough methodology point to the conclusion that behavioural treatments are superior in a way that cannot be explained by the relative methodological quality of the research procedure. The ability of non-behavioural treatments to generalise gains from therapy research to clinical practice is examined in a paper published in 1997 (Shadish et al. 1997). The tentative conclusion is that therapy conducted under clinically representative conditions has comparable effect sizes to those yielded by past meta-analyses. These results are far from definitive due to the uncertainty over what constitutes clinical representativeness and lack of alternative methodologies with which to study generalisability. However, the evidence supports the quantitative reviews of therapy outcome studies in their conclusion that study

participants who receive therapy have better outcomes than those who receive none, and further that these findings from outcome studies are likely to generalise to clinical situations.

1.5.4: Key studies

Fonagy makes a good summary of the conditions of critical acceptance when he states the criteria for inclusion in his review of what works for whom in child and adolescent mental health: 'to be included, reports of research had to satisfy criteria of relevance, outcome and design' (Fonagy et al 2002 p37). The criteria were interpreted in the following way:

1. **Relevance.** Studies that reported evaluations with one or more of the target disorder populations of children and adolescents.
2. **Outcome.** Studies were selected only if they reported outcomes that were either directly related to the symptom disorder (e.g. symptom reduction) or related to intermediary variables. In the latter case the reviewers had independent evidence of an impact on mental health associated with the outcome (e.g. major risk factors such as educational progress in conduct-disordered children) or an impact on mental health was plausible (e.g. family dysfunction).
3. **Design.** The review focused on studies that used an experimental or quasi-experimental study design. There was however a necessary relaxation of normal exclusion criteria because a preliminary exploration of the available evidence indicated that the exclusion of poorly controlled studies would drastically curtail the available database to the point where the meaningfulness and relevance of the review might well be seriously jeopardised.

His team then decided on key selection criteria for acceptance into the review:

1. A clear description of the patient population in the study, either in terms of diagnosis, or in terms of specific problems addressed in the treatment.
2. The study was reported in English
3. The study had a group design or an experimental single-case design
4. There was a systematic effort at the measurement of outcome, including at least pre-and post-test measurement on an objective instrument.

Using this analysis as a guide, the studies encountered in this literature review as having direct relevance to the AGI Study are presented in the table 1.5.4, and the way in which they satisfy methodological criteria is summarised. Discussion of the studies' relative validity is begun below and continued in Chapter 5.

Table 1.5.4: Match between Design Components of Key Studies and those Identified as Desirable

	Kolvin 1981	Cowen 1996	Weiss 1999	CPPRG 2002	Reynolds 2003
School Context	Yes	Yes	Yes	Yes	Yes
Randomised Control	Yes	No	Yes	Yes	Yes
Longitudinal Design	Yes	Yes	No	Yes	Yes
Multiple Respondents	Yes	Yes	Yes	Yes	Yes
Real-World Outcomes	Yes	Yes	No	Yes	Yes
Multiple Intervention Conditions	Yes	No	Yes	Yes	Yes
Time-limited (brief) Intervention	Yes	No	Yes	No	No
Group Interventions	Yes	No	Yes	Yes	Yes
Targeted Prevention	Yes	Yes	Yes	Yes	Yes
Multiple Diagnostic Category	Yes	Yes	Yes	No	Yes
Positive Parametric Data: $p<.05$	Yes	No	Yes	Yes	Yes

Cowen.

The Cowan PMHP system (Cowen, Hightower, Pedro-Carroll, Work, Wyman, & Haffey 1996) has been widely applied and its programmes variously evaluated. At its core, the programme consists of a dyadic relationship between a trained helper and an at-risk child, used to facilitate play opportunities and model relationship-skills during the first year of schooling. Using devised behaviour scales, the programme screens for at-risk children and applies the same measures one year later to inform an 'End-of-Program Conference' to assess overall progress in the programme and their current level of functioning. The evaluation of this massive programme is broad but not deep. The validity and reliability of the referral and assessment tools are inadequately reported and inconclusively established (Hightower et al. 1986; Hightower et al. 1987) and the acknowledged lack of attention to randomisation and control (Cowen, Hightower, Pedro-Carroll, Work, Wyman, & Haffey 1996 p148ff) has only belatedly been addressed and then in a partial form (Duerr 1993). The evaluation is effectively holed beneath the water line but it is worth mentioning that the sheer scale of the venture, that can be mentioned in the same breath as Headstart, resists the consignment of PMHP to the no-man's land

of inconclusively validated studies. It is true that the longitudinal data tends to be of a general type of categorical involvement in designated activities, in the manner of Headstart, but data have been published to demonstrate significant pre-post improvements in self-efficacy ratings and social problem solving (Meller et al. 1994).

Weiss.

In this study (Weiss, Catron, Harris, & Phung 1999), children were screened using teacher-, peer-, and self-report mental health measures covering six domains of psychopathology, and the sample weighted to be representative of an out-patient setting. The sample was banded for disturbance characteristic before being randomly assigned to 'Traditional Child Psychotherapy' involving individuals in sessions that varied in treatment, modality, duration and frequency depending on the assessment of the individual by the clinician but was provided for up to two years. Alternatively, participants were assigned to 'Academic Tutoring' delivered by graduates with training in tutoring. Both groups were assessed using a multi-informant assessment approach, with participant assessed at multiple time points. Standardised teacher, parent and self-reports were compared with sociometric peer data. These data were analysed using a mixed-models approach to hierarchical linear outcome. The details of analysis remain unclear (p87), but yield the following results. Parents and teachers of programme completers reported significant improvement in the children's psychopathology over time for both groups; $F(1,276)=6.21, p<0.05$ and $F(1,276)=4.81, p<0.05$ respectively. The Group vs Time interaction was not significant for any other informants. None of the 20 primary analyses produced significant treatment effects. Both groups improved over time according to teachers, parents and children themselves; however, there was no difference between psychotherapy and curriculum control. The explanation offered by Weisz and colleagues is that traditional child psychotherapy is not effective and further attempts made in the paper to limit the ineffectiveness to the specific set of therapists and techniques are unconvincing. But it could more accurately be said that traditional child psychotherapy is no more effective than small-group teaching. The gains made over time could be attributable to maturation, and the lack of no-intervention controls limits the exclusion of this possibility, but could also be due to a 'human contact factor' not controlled for and not detectable in a whole-class situation. The study sheds no light on the question of whether different treatments approaches work for different anti-social characteristics, with only the parents distinguishing between psychotherapy for anxiety

scores and psychotherapy for aggression scores when compared to controls. There was no evidence that psychotherapy was significantly better as a treatment for one over the other.

Conduct Problems Prevention Research Group (CPPRG)

Highlighted consistently in reviews as an example of a successful multi-component intervention programme for conduct problems, the Fast Track program has a good research base, a large sample and longitudinal data (Conduct Problems Prevention Research Group 2002b). Fast Track is a customised hybrid of intervention components including the universally delivered PATHS classroom curriculum (Kusche & Greenberg 1994). The indicated components focus on child social skills (groups), parent training (groups), and parent child relationship enhancement, and individualised components which involve academic tutoring, pairing of children with prosocial peers, and family support via home visiting. Delivered in four diverse communities throughout the USA, Fast Track began with a large sample of Kindergarten children who showed behavioural indications of risk for conduct disorder, and conducted the overall intervention programme from 1st through 10th Grades with the participating children and families (Conduct Problems Prevention Research Group 2002a). This long term intervention trial has a number of noteworthy features such as theoretically driven intervention components; rigorous implementation of interventions, fidelity assessment, and measurement of procedures; integrity in the presentation and interpretation of results; and attention to issues of community engagement, sustainability and replicability (Prinz 2002). Published results to date have indicated that the intervention package after 3 years of programming produced significant though modest prevention or reduction of conduct problems and associated difficulties based on multiple sources of data. The sample was $n=9000$, with entire schools being allocated to intervention or no-intervention control conditions. The outcome data were collected from teachers, parents and peers relating to conduct, cognition, academic progress, social competence and parenting behaviour. Of nineteen measures seven indicated significant differences between the intervention and control. Of these seven, two indicated an effect size of more than 0.2. These two moderate effect sizes were reported by teachers and parents, who could not have been blind to intervention status. Overall, these gains are modest, and one might look for more return on a lengthy investment. There is no evidence that

the intervention content produced the improvement or the sustained attention from the Fast Track team.

Reynolds.

This study reports on the Chicago Child-Parent Center (CPC) programme (Reynolds & Robertson 2003). The CPC program provides preschool education for low-income children beginning age 3, and a variety of family support services inside and outside the centres aimed at reducing child abuse and neglect (CAN). Although home visitation is provided, most family support activities are directed toward enhancing involvement in children's education at home and in school. Through programme activities in parent resource rooms and in community contexts, parenting skills, vocational skills and social supports are strengthened, which are expected to reduce the likelihood of behaviours leading to CAN. The intervention has a longer duration than most other early interventions because child education and family services are offered in kindergarten to third grade to ease transition issues. The research design charts the social and educational development of a cohort of 1,539 low-income minority children who attended early childhood programmes in 25 sites. 900 children completed the preschool and kindergarten programs and completed the school-age component for 3 years following. The comparison group of 550 children in a quasi-experimental design participated in alternative full-day kindergarten programmes for low-income children emphasising educational enrichment, but were not involved in the preschool programmes. The control group matched the CPC group on age, participation in government funded programmes and family poverty. The children were assessed on any indicated report of abuse or neglect from the courts or department of child and family services (DCFS). Findings conclude that participation in the CPC preschool programme and in the extended programme from preschool to third grade was associated with significantly lower rates of CAN through adolescence for both court and DCFS reports. One distinctive finding of this study was that programme measures are not generally associated with reductions in CAN by age 9. Most of the preventative effects of the preschool participation and extended intervention came during the children's second decade.

Kolvin.

The study conducted by Kolvin (Kolvin, Garside, Nicol, Macmillan, Wolstenholme, & Leitch 1981), contains the most control data of any one study in the literature of school

interventions. The findings reported by the study are invaluable when trying to isolate the effective elements from the ineffective, and although the literature has moved on in many ways since it was published in the early 1980's the principles of comparison remain. Kolvin identified children at-risk of emotional and behavioural problems (which he calls maladjustment) and established sets of group interventions in primary (4-11) and secondary (11-16) schools.

The primary programme compared:

1. At-risk no-intervention controls
2. Parent counselling-teacher consultation, undertaken by social workers.
3. Nurture work, undertaken by classroom assistants.
4. Playgroup regime, undertaken by therapists.

The secondary programme compared:

1. "Maladjusted" no-intervention controls.
2. Parent counselling-teacher consultation, undertaken by social workers
3. Behaviour modification undertaken by a psychologist and teachers
4. Group therapy, undertaken by therapists

Considering first the junior children, playgroup children did significantly better than other groups in terms of overall severity but no significant differences emerged between the interventions for Externalising behaviour. For Internalising behaviour, playgroup children did better than all three other groups at final follow-up. Considering senior children, group therapy and behaviour modification did better than the other 2 groups at final follow-up. On Externalising behaviour, group therapy was superior to all other interventions at midline and at 2 years all interventions were significantly better than the at-risk controls. For Internalising behaviours both group therapy and behaviour modification did significantly better than other conditions at midline and final follow-up.

To Kolvin, "the nature of our screens reflected our concept of maladjustment, and the fact that we were particularly interested in its behavioural and social aspects" (p335) He goes on to say... "Our screen techniques gave rise to a wider spectrum of maladjustment than occurs with a screen based on a single scale and was therefore representative of psychiatric disturbance usually found in clinical practices in a wide range of settings" (p335). The screening tools are used at initial assessment, at a midline follow-up point a number of months after the cessation of intervention, and at a final

follow-up 18 months after the midline assessment and provide the data not only for initial selection of the children but also the data on their reaction to the interventions in this longitudinal design. The screening tools themselves include parent report (Rutter A Scale), teacher report (Rutter Scale B2) (Rutter 1967) and self report (Eysenck Personality Inventory) (Eysenck 1975) all designed to give data on psychological adjustment. In addition the screening battery included an attitude scale administered by self-report (Barker Lunn Children's Attitude Scale) (Barker Lunn 1967), and an inventory of behaviour and temperament administered as a semi-structured interview with parents. Additional data gathered included reading and vocabulary ability comprehension and intelligence tests. The point here is not to critique the validity of the individual scales employed as tools, but to illustrate the multi-dimensional nature of the assessment. Behind this multiplicity lies the assumption, already articulated in the quotes from Kolvin above, that the nature of 'maladjustment' is complex and in order to measure it effects and the effects of treatment on the symptoms of disturbance there is a requirement to employ methods that display an awareness of this complexity. Kolvin and his team are not afraid of the extra effort in developing procedures to deal with the additional information, nor of the possibility of dealing with the conflicting data that will arise from using instruments designed to measure different aspects of disturbance: in this case psychological function, classroom function, academic ability, adaptation in the home environment and the outward expression of adaptation with is attitude.

The nurturing programme run by classroom assistants is broadly similar to the Cowan PMHP in the use of non-specialist provision of a generalised supporting role for the child. What is missing for the Kolvin's assistants is the integration of the input from the assistants into a wider framework of services outside the school, but more particularly working into the school. This integration in the Cowan programme gives status to the intervention and affects the nature of the relationship with the child. In addition, Kolvin's supervising teachers were not volunteers and teacher motivation for incorporating this intervention is acknowledged to be mixed (p155). The results are encouraging considering these two limiting factors (significant difference at final follow-up in favour of nurture groups over maladjusted controls – overall severity): better than parent counselling, but not as convincing as Group Therapy. Group Therapy in the junior school has similarities to the nurture groups in that it supports child-generated interaction and adopts a permissive approach to self-expression through play, but its superiority in terms

of outcome must have an explanation. The expertise and understanding of a trained therapeutic outlook is the obvious explanation: Kolvin trained experienced social workers in the rudiments of play therapy and psychodynamics (p233) in a programme of supervised individual case-load analysis. These workers came from backgrounds of childcare, but not psychotherapy. These volunteers were then provided with weekly supervision sessions and access to "emergency" sessions in-between the weekly sessions. One explanation for the difference in outcome between the nurture groups and the therapy groups would be that; although the relationship between the intervener and children was an essential component of both nurture and therapy groups, the relationships in the latter group were enhanced by the marked superiority of status of intervention and intervener. The access to supervision relevant to the intervention group was superior, and this supervision was on-going, serving to further enhance relationships in a developmental framework (see Kolvin's interpretation of Ginott (1961) in the principles of therapeutic engagement p221ff). According to this analysis the difference in significant outcome favouring group therapy is less to do with difference in form of intervention and theoretical orientation of intervener and more to do with the status of intervener and intervention, alongside the support and training given to developing formative and transformative relationships in the intervention period. If this assumption that the positive effects of intervention are closely allied to the depth of and support for emotionally enriching relationships is applied to the senior school interventions, a more confusing picture emerges. In the senior schools there is parity in results between the behaviour modification programme and group therapy; both are significantly superior to maladjusted controls and parent consultation conditions. These results are problematic if the quality of the relationship is considered as the primary therapeutic force in changing anti-social behaviour in schools. For the guiding principles of behaviour therapy emerge from a background of strictly cause and effect: that behaviour is responsive to particular stimulus configurations, and that behaviour may be modified by its consequences. Traditionally, behaviour is conceptualised as being shaped by the effect of external cues and reinforcers rather than caused by internal drives and forces, and this has made identification of principles for modifying children's behaviour possible. These principles are variously reported (Kruboltz, Rutter) and from a list of 13, only 3 relate to the techniques employed by Kolvin's behaviour modification programme. The 3 principles of Positive Reinforcement, Modelling (imitation) and Cueing /Prompting are amongst the simplest and easily applied principles, yielding

techniques that are non-invasive and easily incorporated into a positive teaching demeanour. This is a pertinent observation because of the acknowledged overlap in behavioural and psychotherapeutic techniques (Howard & Kendall 1996; Meichenbaum 1997), particularly amongst those simpler, non-invasive techniques. It is argued that in the specific case of the interventions delivered to senior schoolchildren in the Kolvin study, the traditional distinction between behavioural and psychotherapeutic approaches is blurred: the behaviour modification treatment condition contains significant elements of relationship building more usually classified as psychotherapeutic. Kolvin's emphasis on the techniques of social reinforcement relying on the systematic and contingent use of teacher praise, attention and approval rather than techniques based upon the use of material or concrete rewards (p107), is an example where behavioural techniques in the strictest sense of the word were considered less favourably than techniques with an overlap with a more therapeutic approach. Teacher praise and attention cannot be spoken of in the same way as a material reward, for with praise and attention come human warmth and regard, the bare bones of relationship. Added to this, the majority of cases dealt with in the Kolvin study were one or two children in a class of thirty. A special focus of praise and attention on two children out of thirty will have effects far beyond those of stricter cause-and-effect behavioural techniques. The child will feel that contrary to previous experience, his/her efforts are noticed and praised, she will begin to feel more integrated into the class where previously she felt only hopelessness. More importantly the change in the relationship with the teacher will be significant; where before the teacher will have expressed frustration or a deafening silence when considering progress, now there will be warmth and attention. It is difficult to avoid the conclusion that the Behaviour Modification programme tested by Kolvin may have been closer in content to the Group Therapy programme than at first thought.

Section 1.6: Trigger for AGI Research Project

The triggers for this research are the gaps that analysis and experience have revealed. Gaps in practice, research and policy that have been encountered in that order. By attempting to fill the gaps it is the hope of the research team that positive solutions will be generated for all those concerned in youth disaffection. The gaps in practice are in the scarcity of community-based programmes staffed to meet the scale of disaffection,

based on theory that accounts for the complexity of disaffection delivering outcomes that are positive for all those involved in disaffection. The gaps in research are in the scarcity of sensitively evaluated community-based programmes of the type outlined above, evaluations that are rigorous, analytical, empowering persuasive to multiple audiences and readable. The gaps in (UK) policy lie in the lack of cohesion initiatives display: initiatives designed to facilitate intervention display piecemeal and partial analysis and confusion in programme design. This research project is primarily concerned with the gaps in research, but the trigger for the research must be seen in how practice influences research and how research influences policy. The socially engaged aspect of the research is an integral part of the context in which it should be considered. Through this contribution to filling the gaps in research, it is hoped that influence is brought to bear on policy, in order that the provision for those disadvantaged by circumstance or personal outlook may be better cared for. Needless to say, the influence should also be seen to flow backwards to inform the practice out of which the research springs, by providing new frameworks to understand the behaviours encountered, and through application of the insights from evaluation.

The discussion contained in previous sections made the case for criteria that an intervention must satisfy in order to be considered effective. The criteria were grouped under the headings of scale, complexity and outcome. The studies considered in section five all satisfied stated criteria of being group interventions conducted in community settings delivering an intervention based on a pan-theoretical understanding of youth function. The studies adhere to the conventions of methodological discipline and demonstrate outcomes that impact in real-world settings. However the analysis above has identified areas where the literature is thin and present studies do not deliver fully. These are the specific triggers for the research contained here and they fall into the categories of scale, complexity and outcome.

Scale: there has been little consideration of the impact of group interventions delivered in a school setting. In addition while the literature is rich in the 'what works for whom?' type of research that seeks the grail of treatment matched to symptom, but there is less evidence of which interventions work for the heterogeneity of symptom and severity found in community samples.

Complexity: As a result of the tendency towards specificity in the research literature the underlying pan-theoretical and developmental understandings of disaffection are often

implicit (see above). Studies of explicitly developmental interventions represent the cutting edge of outcome research.

Outcome: research that uses multiple respondents is becoming more common, but there is a gap in systematic consideration of multiple respondents to inform whether outcomes are respondent specific.

Partly in response to the situation described above, The Learning Challenge was set up to work in schools with children and staff on developing therapeutic behaviour management practices. This research project has been designed, delivered and written up in keeping with the vision of TLC that attempts to articulate what is possible for children and schools. This vision, and the action necessary to achieve the vision are driving forces in developing the arguments contained here. The statement has been through stages of development, but in the most recent literature is given as:

'The Learning Challenge sees children as achieving their potential when they are working towards being self-motivated learners able to negotiate complex peer and hierarchical relationships with an increasing element of self-awareness. Furthermore children will be achieving their potential when they gain a growing insight into their own internal processes in dealing with family relationships and they use their emotional health to contribute to long-term goals that include a wider community.' (The Learning Challenge. 2004 p1)

Because child disaffection does not take place in a vacuum the importance of the research is thought to cover the social context, and to have meaning for those involved in the social context. Explicitly stated this means:

'The Learning Challenge believes that this research is important because the methods of treating the symptoms of disaffection in education, health and social service contexts are seldom of a sufficient scale to match the extent of need. Where preventative initiatives exist in the community contexts most appropriate for interventions, there is a need for systematic evaluation of a type persuasive for policy-makers. In highlighting a group intervention and providing data to allow policy-makers to make decision about its efficacy, we hope that future statutory service provision will reflect evidence-based preventative practice. Furthermore we believe that a service will have been done for

deprived, disaffected and unhappy children in the process.' (The Learning Challenge. 2004 p1)

Section 1.7: Definition of AGI Research Project

1.7.1: Aims

Grand Aims: the AGI study aims to provide data helpful to service delivery in school, helpful to practitioners dealing with the challenge of disaffection, and to the children who are themselves disaffected. In addition, the aim was to provide a basis for policy in planning and developing effective interventions in youth disaffection in the school setting.

Pragmatic Aims: the AGI study aims to develop interventions capable of delivering change in disaffection with lasting effects over time. Interventions capable of working with schools in meeting the challenges of disaffection and building relationships with staff teams and parents.

Research Aims: concerning the development of a methodology able to provide data that was relevant to the intervention and reliable in measuring any change over time. To reflect on the findings with depth and in relation to other studies of similar aim with the purpose of finding meaning in positive and negative results. This meaning was intended to feed back into the cycle of policy and practice of the grand aims.

It is considered important that the research is conducted in keeping with the above statement and does nothing to intrinsically damage children, but does assist children working towards the ambitious aims stated there. It is therefore most important that the research does not exploit any of the vulnerable parties asked to take part, but it is equally important that the research is conducted, written and published in such a way so as to draw positive attention to children involved in achieving this aim.

All the sources from which data are drawn are important, but in particular the children themselves should feel a sense of self-determination through their involvement with the research process, indeed wherever possible the process itself should assist them in achievement of the aims stated above. The priorities remain with the children working in

the research, their welfare and their continued development. In addition, the research audience is a priority, in order that the positive attention sought is achieved. We believe that it is important that any positive outcome from the research, whether in terms of recognition, or increased resources, should be focussed on the communities providing the data for the research; schools, children, parents and so on.

1.7.2: Definition

This research project presents data on the comparative effectiveness of two group interventions delivered in the school setting to children aged between 11 and 15 years old. The interventions differed only in their content: one being based on a model of developmental psychopathology, the other using child centred teaching methods to encourage learning skills acquisition. Data were gathered from multiple sources over a period of 1 year and outcomes compared.

In order to for the wider social aspects of the research to be satisfied (these form part of the trigger for research), methodological rigour was a priority, with 133 children randomised to the intervention conditions and a longitudinal quasi-experimental design initiated with the aim of providing data capable of impacting on policy.

The research project has set out to make particular contributions in the following areas that correspond to the considerations of scale complexity and outcome and are designed to expand the existing published literature:

The group interventions are non-specific in terms of their treatment framework and the symptomology of the client group. The effectiveness of such interventions is under-represented in the literature, but of increasing importance as developmental theory becomes the dominant interpretive mode in youth disaffection (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002).

The interventions are designed to be effective in the school context, with a school population. A trained schoolteacher delivers the intervention. This is an addition to the literature in (UK-based) school-studies.

The interventions are set up so as to compare the effectiveness of a developmental intervention framework with a child centred skills acquisition model. This specific comparative data are absent from the literature.

The form of the outcome data (questionnaire, observation and 'real-world') and the multiple respondents included in the study allow conclusions to be drawn about outcome that account for the practical and philosophical issues in demonstrating outcome.

Section 1.8: Importance of AGI Research Project

Unless effective solutions are found to the phenomenon of youth disaffection that are effective in dealing with the large numbers involved and effective across the exterior contexts of home school and community, as well as the interior contexts of emotions, behaviours and social interaction then treatments will always be poor. A poor use of resources, a poor offering to young people disadvantaged by circumstance, a poor appreciation of our chalk-face public servants and a poor reflection on our sophistication as a society.

This research project should be seen in the context of a cycle of expertise between practice, research and policy. The immediate contribution is to the literature of school-based community interventions with disaffected youth, but it is hoped that the rigour and accessibility of the methodology together with the emphasis on transferability and real-world outcomes makes this a contribution to practitioners and policy-makers.

Reflective Commentary

These sections, and the sections like it throughout this thesis, are aimed at providing a view of the preceding chapter as a coherent whole and addressing issues of reflexivity and bias. Through use of narrative, it also provides the opportunity to contextualise research writings in the real-world situation where it took place. Chapter One takes a look at the work that has gone on in theorising and researching youth disturbance and disaffection, providing a comprehensive review of published studies and papers relating to interventions. Although the result is regarded as well structured and thorough, this section serves as a reminder of the agenda towards which conclusions have been pulled.

A recurring theme that will emerge as this account of the AGI study goes on is how the research emerged out of practice newly established in the North Tyneside region in the North East of England. As a schoolteacher and arts-therapist working with disaffected children, Toby Quibell worked hard to find a form of intervention that was within his competence while providing emotional and behavioural outcomes of benefit within the school setting. This turned out to be group-based sessions delivering social and emotional skills development within a creative-expressive environment (the AGI is detailed elsewhere). It seems clear that the literature has been evaluated using clear critical guidelines to distinguish quality, and conclusions built on trends that have been discerned as a result. Caution needs to be exercised when filleting out the group approaches under-represented in the published literature, especially where studies contain similarities with the developed practice. The balanced position found in the discussion about behavioural and cognitive behavioural interventions on page 44 is an example of how this cautionary approach can strengthen the critical case for alternative treatment approaches and allows those areas of writing less favoured by publication to be drawn out and established. The drive of the critique offered in the preceding sections does serve to raise the profile and enhance the credibility of research echoing the elements contained in AGI practice, but balanced assessment of the presentation of the argument suggests this is justified by standards of critical evaluation.

In this vein, it is noted on page 24 that “some sort of outcome is usually present for most treatments” and this advanced as a rationale for looking beyond the initial presentation of positive findings to the areas of weakness and methodological uncertainty. This is also true of the AGI study and these pages can be seen as a platform for the achievements of the research. Where these claims and interpretations are optimistic and over-stated the same rigour will need to be applied to clarify the way negative findings are presented with a positive implication for the AGI programme.

The discussions on prevalence of disturbance and disaffection provide an interesting reflection on the experience of working in the primary middle and secondary schools of North Tyneside. Those schools selected are situated in areas of multiple deprivation, and up to half of the class population was included from STC and MWP primary schools. This stretches the upper limit of disaffection of 25%, but was felt to work in both cases. The mixture and composition of the group necessitated the inclusion of children not

considered disaffected, but in both these schools it is no exaggeration to consider all class members as at-risk of disaffection and it is more remarkable to reflect on the numbers of children that do go on from these beginnings to fulfil some of their potential in healthy interaction with their communities. In JSP School, while running a session as part of this study, there was a disturbance outside the classroom when the outer wall joined the playground. The group became fascinated with the plumes of smoke periodically appearing from beneath the window and waited expectantly for the facilitator to do something. How much more interesting to be a spectator in the process of mutual humiliation that is high-level teacher-pupil discipline than to engage in groupwork. The facilitator approached the window with a sense of foreboding, enhanced upon the sight of a notorious (tiny) Year 7 boy who was considered uncontrollable in behaviour or language. The group in the classroom were metaphorically rubbing their hands with glee as he tapped on the window. The boy looked up and without a moment's hesitation and with a rather wonderful disregard for the consequences stuck up two fingers and said "Fuck off you ginger bastard". Pretty it wasn't, but it was compact and utilitarian in its core message. These schools were not picked for their likelihood of providing successful outcome, but rather for the level of need in their classrooms.

Chapter Two: Questions & Hypotheses

Introduction

The research team conducted this study into group interventions in youth disaffection because driven by a desire to discover whether interventions that meet the scale, complexity and need for outcome can be effective. A clear aim has been to understand the nature of disaffection and provide evidence upon which more effective public services can be built. These questions are set out to give structure to the enquiry and a deliberate sense of scientific rigour to the nature of the data collected. The experience of the research team in their encounters with disaffection found much that is chaotic about the behaviours of young people along with a tendency to make generalisations about the effectiveness of treatment approaches. Upon entering MWP school at the time of the interventions, the overwhelming impression was of stepping back in time: doubtless there were clean children whose parents struggled against the odds to keep them stable and presentable, but the everywhere the eye would roam there were children whose behaviour and appearance seemed locked in a previous decade. In this primary school the children, for all their bravado and swagger were curiously wary and respectful of the female staff. Possibly this was due to the social make-up of the estate, where generations of female family members lived in close proximity, frequently without male partners. The dress-code seemed to ignore most fashion dictates of the time, perhaps for economic reasons, but more because the estate seemed to operate its own micro-climate with regard for social norms, and just how much disaffection was to be tolerated. Rumour had it that the two opposing families 'running' the estate had agreed to exclude drugs and firearms in their black-market activities. Whatever the underlying reasons, the result was a curiously innocent toughness, a kind of honest-to-goodness badness that encouraged a motherly approach from most of the professionals involved in their education. For Toby Quibell working exclusively with the anti-social fall-out from the classrooms, the clarity of a quantitative study with comparable groups receiving slightly different interventions seemed to hold the key to cutting through the chaos of behaviours and the mirror-image of that chaos in the prescriptions made by school teachers and their assistants. The key question for the research is "Can these group interventions demonstrate a positive effect on a range of behaviours over a range of contexts?" By comparing two groups of children going through different group experiences and

examining the impact on aspects of emotional and social behaviour from different sources the aim was to answer the core research questions below. These questions were developed out of the experience of working with troubled children on the Meadow Well estate in North Shields and reflect a practitioner's concern with meeting the challenge of anti-social behaviours.

Section 2.1: Questions

1. Does working in a group have an effect on a child's emotional/behavioural profile?

Do children gain in the skills reported by questionnaires when they are involved in group interventions?

1(sub). Are any gains in emotional/behavioural profile reported during the group intervention greater than those over a similar period when no intervention is taking place?

If you compare the time when children have been working in a group with the time when they have not, will you notice a difference in their behavioural and emotional ratings?

2. Do gains in emotional/behavioural profile made while the groups are running last after the group is finished?

If you survey the respondents at intervals after the interventions are over will you still be able to detect the gains made?

3. Do gains in emotional/behavioural profile depend on what goes on in the group?

If you compare a group who are engaged in child-focussed developmental and therapeutic activities (AGI) with a group engaged in curriculum tasks (CGI), will you notice a difference in emotional/behavioural profile?

4. How do any differences in emotional/behavioural profile between the groups alter as time goes by?

If you survey the respondents at intervals after the interventions are over will you still be able to detect differences in the gains made by the different groups?

5. Does working in a group have an effect on a child's classroom performance and school attendance?

Do children gain in 'real-world' skills when they are involved in group interventions?

5(sub). Are any effects on a child's classroom performance and school attendance reported during the group intervention greater than those over a similar period when no intervention is taking place?

If you compare children who have been working in a group with those who have not, will you notice a difference in their 'real-world' skills?

6. Do the effects on a child's classroom performance and school attendance made while the groups are running last after the group is finished?

If you survey the population at intervals after the interventions are over will you still be able to detect the gains made?

7. Do the effects on a child's classroom performance and school attendance depend on what goes on in the group?

If you compare a group who are engaged in child-focussed developmental and therapeutic activities (AGI) with a group engaged in curriculum tasks (CGI), will you notice a difference in 'real-world' skills?

8. How do any differences in children's classroom performance and school attendance between the groups alter as time goes by?

If you survey the respondents at intervals after the interventions are over will you still be able to detect differences in the 'real-world' skills held by the different groups?

9. Does the attitude of the school to the intervention and research process affect the outcomes?

If you allocate a score to the schools involved based on a number of criteria relating to the understanding of and interest in the intervention and AGI research programme will you notice those with higher scores having most of the best outcomes?

Section 2.2: Summary of Design

130 children were randomised to one of two group intervention conditions. Data were gathered using instruments completed by teachers, parents and children themselves. In addition a direct observation protocol was used along with attendance and achievement data from school records. The instrument battery was chosen to give data relating to a range of emotional and behavioural states assessed in a range of contexts. Data was collected at four time-points.

T1 – three months prior to intervention

T2 – immediately prior to intervention

T3 – immediately following intervention

T4 – one year following the cessation of intervention.

Questionnaires were used to gather the data relevant to Questions 1, 2, 3 & 4.

Two self-reports are used: the Youth Self Report (YSR) (Achenbach 1991) and the Multi-Dimensional Self Concept Scale (MSCS) (Bracken 1993).

A teacher report is used: the Teacher Report Form (TRF) (Achenbach 1991).

A parent report is used: the Child Behaviour Checklist (CBC) (Achenbach 1991).

Independent observation and school attendance reports were used to gather the data relevant to Questions 5, 6, 7 & 8.

An on-task observation protocol is used for group members.

A whole-class off-task observation is used to compare the group to the rest of the class.

Attendance figures are used to gauge the wider school impact.

Intervention One

The Action Groupskills Intervention (AGI) uses an eclectic approach designed to connect with a child's experience and explore that experience in a contained group environment.

Designed to maximise the impact of a positive relationship network that includes the facilitator.

Intervention Two. A small Curriculum Studies Group (CSG) uses focussed teacher attention, praise encouragement and rewards based on performance and progress in

curriculum tasks undertaken in the group setting. The detail about the interventions is found in Chapter 3, Section 2.

Section 2.3: From Questions to Hypotheses

Question 1. Does working in a group effect a child's emotional/behavioural profile?

Do children gain in the skills assessed by questionnaires when they are involved in group interventions?

Question 1(sub). Are any gains in emotional/behavioural profile reported during the group intervention greater than those over a similar period when no intervention is taking place?

If you compare the time when children have been working in a group with the time when they have not, will you notice a difference in their behavioural and emotional ratings?

What is the effect of working in a small group compared to usual school class experience? To answer this, both the AGI and the CSG groups are considered together, and time points T1, T2 and T3 are compared. It is predicted that positive feedback, group purpose and forces of group identity will have an effect of reducing problem behaviours and encouraging self-esteem. Hypothesis No.1 therefore states that it is expected that these interventions will have positive effects over the whole period (T1 – T3). As this period contains a waiting-list period (T1 – T2) during which there was no intervention, the sub hypothesis is that the gains during this T1 – T2 period will be less than during the T2 – T3 intervention period.

Hypothesis 1:

According to self-report, teachers and parents, working in groups reduces behavioural and emotional problems and promotes self-esteem as reported from initial assessment to post-intervention (T1 – T3).

Hypothesis 1 (sub):

According to self-report, teachers and parents, the gains over the no-intervention period (T1 – T2) will be less than over the subsequent intervention period (T2 – T3).

Question 2. Do gains in emotional/behavioural profile made while the groups are running last after the group is finished?

If you survey the respondents at one year after the interventions are over will you still be able to detect the gains made?

Are the effects of working in a group sustained after intervention has finished? To answer this, both AGI and CSG groups are considered together, and time points T2, T3 and T4 are considered. The input common to both AGI and CSG groups is intended to have a lasting impact in the way a child behaves and based on the reading of the literature in Chapter One, this effect should be present one year after intervention. It is therefore predicted that any improvement for the intervention groups over the intervention period T2-T3 will be sustained one year later at T4. The hypothesis looks for significant improvement over intervention and no significant change over the follow-up period.

Hypothesis 2:

According to self-report, teachers and parents, gains made on emotional, behavioural and self-esteem measures over the intervention period (T2 – T3) will be sustained over the follow-up period (T3 – T4).

Question 3. Do gains in emotional/behavioural profile depend on what goes on in the group?

If you compare a group who are engaged in child-focussed developmental and therapeutic activities (AGI) with a group engaged in curriculum tasks (CSG), will you notice a difference in emotional/behavioural profile?

Is there any difference between the groups in the short-term? To answer this the impact on the AGI group can be compared to the impact on the CSG group the intervention period T2-T3. The groups are experimentally controlled for variables such as time of day, duration, facilitator and general child-centred approach, but are different in the content of the sessions delivered. Because the AGI sessions are more clearly focussed on the enrichment of relationships and the resolution of conflict situations, it is predicted

that the AGI group will be more effective in reducing problem behaviours and increasing self-esteem.

Hypothesis 3:

According to self-report, teachers and parents, the groupskills intervention (AGI) will be more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems and promoting self-esteem.

Question 4. How do any between-group differences in emotional/behavioural profile between the groups alter as time goes by?

If you survey the respondents at one year after the interventions are over will you still be able to detect differences in the gains made by the different groups?

Is it possible to distinguish between the intervention groups at a follow-up point of one year? To answer this the AGI and CSG groups are compared at one year following intervention to see if there is any difference between the scores on the questionnaires. The AGI intervention is considered to be more able to affect the multiple contexts in which an individual will express disaffection also the multiple areas of emotional and social function that need to be addressed in order to provide lasting change. It is therefore predicted that the positive impact on AGI intervention group will be greater at the follow-up point of one year (T4) than the CSG group whose intervention has focussed on using the group to achieve curriculum tasks.

Hypothesis 4:

According to self-report, teachers and parents, the AGI group will show greater long-term benefits on behavioural, emotional and self-concept measures than the CSG group, measured prior to intervention and at one year.

Question 5. Does working in a group have an effect on a child's classroom performance and school attendance?

Do children gain in 'real-world' skills when they are involved in group interventions?

Question 5(sub). Are any effects on a child's classroom performance and school attendance reported during the 12 weeks of group intervention greater than those over a similar period when no intervention is taking place?

If you compare children when they have been working in a group over 12 weeks with the same children over a similar period when they have not, will you notice a difference in their 'real-world' skills?

What is the effect of working in a small group compared to the usual classroom experience? In order to answer this question the AGI group and the CSG group are considered together over the whole period T1-T3, from the beginning of the waiting list period to the end of intervention. It is predicted that positive feedback, group purpose and forces of group identity will have an effect of reducing 'real-world' problem behaviours observable as on- and off- task behaviours and attendance rates. Hypothesis No.5 therefore predicts that these interventions will have positive effects over whole period (T1 – T3). As this period contains a waiting-list period (T1 – T2) during which there was no intervention, the sub hypothesis is that the gains during this T1 – T2 period will be less than the T2 – T3 intervention period.

Hypothesis No.5:

According to independent observation and school attendance records, the intervention group (AGI+CSG) shows gains in classroom behaviour and school attendance reported from initial assessment to post-intervention (T1 – T3).

Hypothesis No.5(sub):

According to independent observation and school attendance records, improvements in the intervention group (AGI+CSG) over the no-intervention period (T1 – T2) will be less than over the subsequent intervention period (T2 – T3).

Question 6. Do the effects on a child's classroom performance and school attendance made while the groups are running last after the group is finished?

If you survey those involved in groupwork at intervals after the interventions are over will you still be able to detect any impact made in 'real-world' behaviours?

Are the 'real-world' effects of working in a group sustained after intervention has finished? In order to answer this question both AGI and CSG groups are considered together and time points T2, T3 and T4 are considered. The input common to both AGI and CSG groups is intended to have a lasting impact in the way a child behaves and based on the reading of the literature in Chapter One, this effect in 'real-world' settings should be present one year after intervention. It is therefore predicted that any improvement for the intervention groups over the intervention period T2-T3 will be sustained one year later at T4. The hypothesis looks for significant improvement over intervention and no significant change over the follow-up period.

Hypothesis 6:

According to independent observation and school attendance records, gains made in classroom behaviour and school attendance over the Intervention period (T2-T3) are sustained over the follow-up period (T3-T4).

Question 7. Do the effects on a child's classroom performance and school attendance depend on what goes on in the group?

If you compare a group who are engaged in child-focussed developmental and therapeutic activities (AGI) with a group engaged in curriculum tasks (CGI), will you notice a difference in 'real-world' skills?

Is there any difference between the groups in the short-term? To answer this the impact on the AGI group can be compared to the impact on the CSG group the intervention period T2-T3. The groups are experimentally controlled for variables such as time of day, duration, facilitator and general child-centred approach, but are different in the content of the sessions delivered. Because the AGI sessions are more clearly focussed on the enrichment of relationships and the group encouraged to consider positive management of situations outside the group sessions, it is predicted that the AGI group will be more effective in reducing problem behaviours and increasing self-esteem.

Hypothesis 7:

According to independent observation and school attendance records, the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.

Question 8. How do any differences in children's classroom performance and school attendance between the groups alter as time goes by?

If you survey the respondents at intervals after the interventions are over will you still be able to detect differences in the 'real-world' skills held by the different groups?

Is it possible to distinguish between the intervention groups at a follow-up point of one year? To answer this the AGI and CSG groups are compared at one year following Intervention to see if there is any difference between the non-questionnaire data. The AGI intervention is considered to be more able to affect the multiple contexts in which an individual will express disaffection also the multiple areas of emotional and social function that need to be addressed in order to provide lasting change. It is therefore predicted that the positive impact on AGI intervention group will be greater at the follow-up point of one year (T4) than the CSG group whose intervention has focussed on using the group to achieve curriculum tasks.

Hypothesis 8:

According to independent observation and school attendance records, there will be a measurable difference between gains in classroom behaviour and school attendance resulting from the groupskills intervention (AGI) and those gains promoted by the curriculum studies intervention (CSG) at a follow-up point of one year.

Question 9. Does the attitude of the school to the Intervention and research process affect the outcomes?

If you allocate a score to the schools involved based on a number of criteria relating to the understanding of and interest in the Intervention and AGI research programme will you notice those with higher scores having most of the best outcomes? Will one intervention condition be more sensitive to these conditions than the other?

The emphasis given to group factors and school factors in section 3.1.1 draws out their importance in establishing group norms for behaviour and the consistency of condition necessary to develop themes of study over a number of weeks. The support of the teaching staff is considered to be vital in contextualising the interventions and a sympathetic policy with regard to the provision of facilities is thought to be vital to the generation of positive behaviour outcomes. For these reasons, the prediction is that a positive attitude of the school hierarchy, and a facilitating approach to establishing conditions within the school will have a positive effect on behaviour.

Hypothesis 9:

Positive behaviour outcome for both AGI and CSG groups will correlate significantly with the attitude of the school to the interventions and research process, measured by scoring within-school conditions.

Section 2.4: Hypotheses

Hypothesis 1

According to self-report, teachers and parents, working in groups reduces behavioural and emotional problems and promotes self-esteem as reported from initial assessment to post-intervention (T1 – T3).

The Hypothesis is split into four parts (a – d) each relating to a different dataset.

- a) According to child self-report (YSR), working in groups reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).
- b) According to child self-report (MSCS), working in groups regard promotes self-esteem as reported from initial assessment to post-intervention (T1 – T3).
- c) According to teacher report (TRF), working in groups reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).
- d) According to parent report (CBC), working in groups reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).

Hypothesis 1 (sub)

According to self-report, teachers and parents, the gains over the no-intervention period (T1 – T2) will be less than over the subsequent intervention period (T2 – T3).

The Hypothesis is split into four parts (a – d) each relating to a different dataset.

- a) According to child self-report (YSR), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).
- b) According to child self-report (MSCS), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).
- c) According to teacher report (TRF), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).
- d) According to parent report (CBC), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).

Hypothesis 2

According to self-report, teachers and parents, gains made on emotional, behavioural and self-esteem measures over the intervention period (T2 – T3) will be sustained over the follow-up period (T3 – T4).

The Hypothesis is split into four parts (a – d) each relating to a different dataset.

- a) According to child self-report (YSR), gains made on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).
- b) According to child self-report (MSCS), the gains on self-esteem measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).
- c) According to teacher report (TRF), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).

- d) According to parent report (CBC), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).

Hypothesis 3

According to self-report, teachers and parents, the groupskills intervention (AGI) will be more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems and promoting self-esteem.

The Hypothesis is split into four parts (a – d) each relating to a different dataset.

- a) According to child self-report (YSR), the groupskills intervention (AGI) will be more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 – T3).
- b) According to child self-report (MSCS), the groupskills intervention (AGI) will be more effective than the small-group curriculum studies intervention (CSG) in promoting self-esteem over the intervention period (T2 – T3).
- c) According to teacher report (TRF), the groupskills intervention (AGI) will be more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 – T3).
- d) According to parent report (CBC), the groupskills intervention (AGI) will be more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 – T3).

Hypothesis 4

According to self-report, teachers and parents, the AGI group will show greater long-term benefits on behavioural, emotional and self-concept measures than the CSG group, measured prior to intervention and at one year.

The Hypothesis is split into four parts (a – d) each relating to a different dataset.

- a) According to child self-report (YSR), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.
- b) According to child self-report (MSCS), the AGI group will show greater long-term benefits on self-concept measures than the CSG group, measured prior to intervention and at one year.
- c) According to teacher report (TRF), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.
- d) According to parent report (CBC), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.

Hypothesis 5

According to independent observation and school attendance records, the intervention group (AGI+CSG) shows gains in classroom behaviour and school attendance reported from initial assessment to post-intervention (T1 – T3).

The Hypothesis is split into three parts (a – c) each relating to a different dataset,

- a) According to independent observation, the intervention group (AGI+CSG) increases on-task classroom behaviour reported from initial assessment to post-intervention (T1 – T3).
- b) According to independent observation, the intervention group (AGI+CSG) shows a reduction in off-task behaviours when compared to the rest of the class, as reported from initial assessment to post-intervention (T1 – T3).
- c) According to school attendance records, the intervention group (AGI+CSG) shows increased attendance when reported from initial assessment to post-intervention (T1 – T3).

Hypothesis 5(sub)

According to independent observation and school attendance records, improvements over the no-intervention period (T1 – T2) will be less than over the subsequent intervention period (T2 – T3).

The Hypothesis is split into three parts (a – c) each relating to a different dataset,

- a) According to independent observation, the intervention group (AGI+CSG) on-task classroom behaviour gains over the intervention period (T2-T3) will be greater than those made in the no-intervention period (T1-T2).
- b) According to independent observation, intervention group (AGI+CSG) shows a greater reduction in off-task classroom behaviour compared to the rest of the class over the intervention period (T2-T3) than those made in the no-intervention period (T1-T2).
- c) According to school attendance records, the intervention group (AGI+CSG) shows a greater increase in attendance over the intervention period (T2-T3) than in the no-intervention period (T1-T2).

Hypothesis 6

According to independent observation and school attendance records, gains made by the intervention group (AGI+CSG) in classroom behaviour and school attendance over the intervention period (T2-T3) are sustained over the follow-up period (T3–T4).

The Hypothesis is split into three parts (a – c) each relating to a different dataset,

- a) According to independent observation, gains made by the intervention group (AGI+CSG) in on-task classroom behaviour over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).
- b) According to independent observation of the intervention group (AGI+CSG), (AGI+CSG): rest of class off-task behaviour ratio reduces over the intervention period (T2 – T3) and is sustained over the follow-up period (T3 – T4).

- c) According to school attendance records, the gains by the intervention group (AGI+CSG) in percentage attendance over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).

Hypothesis 7

According to independent observation and school attendance records over the intervention period, there will be a measurable difference between AGI and CSG groups, with scores resulting from the devised groupwork intervention (AGI) exceeding those scores promoted by the curriculum studies Intervention (CSG)

The Hypothesis is split into three parts (a – c) each relating to a different dataset,

- a) According to Independent observation over the intervention period, there will be a measurable difference in on-task behaviours between AGI and CSG groups, with scores resulting from the devised groupwork intervention (AGI) exceeding those scores promoted by the curriculum studies intervention (CSG)
- b) According to independent observation over the intervention period, there will be a measurable difference in off-task behaviours ratios between AGI and CSG groups, with AGI: rest of class off-task behaviour ratio scores lower than CSG: rest of class off-task behaviour ratio scores.
- c) According to school attendance records over the intervention period, there will be a measurable difference between AGI and CSG groups with percentage attendance scores from the AGI group higher than percentage attendance scores from the CSG group.

Hypothesis 8

According to independent observation and school attendance records at follow-up, there will be a measurable difference between AGI and CSG groups, with scores resulting from the devised groupwork intervention (AGI) exceeding those scores promoted by the curriculum studies intervention (CSG) at a follow-up point of one year.

The Hypothesis is split into three parts (a – c) each relating to a different dataset.

- a) According to independent observation at follow-up, there will be a measurable difference between AGI and CSG groups, with scores made in on-task classroom behaviour resulting from the groupwork intervention (AGI) exceeding those scores promoted by the curriculum studies intervention (CSG) at the follow-up point of one year.
- b) According to independent observation at follow-up, there will be a measurable difference between AGI and CSG groups, with AGI: [rest of class] off-task behaviour ratio scores lower than CSG: [rest of class] off-task behaviour ratio scores at the follow-up point of one year.
- c) According to school attendance records at follow-up, there will be a measurable difference between AGI and CSG groups, with in percentage attendance by the AGI group exceeding percentage attendance made by the CSG group at a follow-up point of one year.

Hypothesis 9:

Positive behaviour outcome for both AGI and CSG groups will correlate significantly with the attitude of the school to the interventions and research process, measured by scoring within-school conditions.

The Hypothesis is split into three parts (a – c) each relating to a different dataset.

- a) When schools are ranked for outcome using a gain score (T2-T4) from child self-report (YSR), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions
- b) When schools are ranked for outcome using a gain score (T2-T4) from teacher report (TRF), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions
- c) When schools are ranked for outcome using a gain score (T2-T4) from parent report (CBC), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions

Reflective Commentary

This chapter sets out the stall when it comes to the approach of the research team to the area of study. The progression from core questions to research questions to hypotheses is presented in the context of a practitioner's response to behaviour occurring in actual schools. The presentation itself is handled as if it were the only viable response to the situation as it was being experienced. On reflection it can be seen as one response from set of choices that could have been used to guide the enquiry. As it stands, the chapter does convey a sense of clarity about what is sought from the data and the way the data will be gathered. By implication also, the nature of the analysis is hinted at – for we are looking for explanations of any difference in outcome between the groups. As a whole therefore the chapter communicates a great deal about the study and the view of its subjects. The hard line scientific approach was favoured as much for its perceived persuasive power in the academic world as for its power to make practice more easily justifiable. The choice to make the primary data questionnaire-based and subject to quantitative analysis necessarily limits the interpretations of behaviours. Investigating the process through which change occurs is a casualty of this choice, as is the impact felt in the community. The inclusion of the non-questionnaire data, despite confirming the bias towards the quantitative, can be seen as an attempt to redress the balance somewhat. This data, with its focus on school attendance and classroom behaviour goes some way towards assessing the impact of the interventions in the social context.

Chapter Three: Intervention

Introduction

Interventions offered in this study comprise the Action GroupSkills Intervention (AGI), a focussed use of therapeutic group factors guided to the acquisition of personal and social skills linked with reflective discussion, and a curriculum studies intervention (CSG) designed to use the small group to boost academic performance.

The first and possibly the defining quality of these interventions is their use of the group as the organizing structure for delivering content. Group factors and the benefits of group membership are understood here in the context of the literature of groupwork. Organizing a useful selection from the vast body of material that can be seen as the 'literature of groupwork' is a challenge. Aside from the obvious problem of scale, there are issues around categorizing material, quality (many texts are repeats of a basic how-to-do-it formula), and purpose. To make life easier I have adapted a framework used by Kenneth E. Reid in his helpful study of the use of groups in social work (Reid 1981), and added in a more therapeutic strand and a category to include the tradition of educational groupwork.

Group Therapy, T-Groups and Encounter Groups - this category looks to the group as a key element in the therapeutic process - and that draws heavily upon central traditions of practice within psychotherapy e.g. psychoanalytic, Gestalt, cognitive-behavioural. Allied to this is that whole body of material around family therapy: the standard works would include those by Wilfred Bion (Bion 1961) and Irvin D. Yalom (Yalom 1975). Another tradition of practice that could be said to fall in this strand is that of Training groups (T-groups). Here Lewin's interest in using small groups as training laboratories for teaching people interpersonal skills (Lewin 1948), and the later development of sensitivity-training or encounter groups (Rogers 1970), are examples of the use of groups for interpersonal learning.

Case-Focused Groupwork - this approach can be described as 'preventative and rehabilitative', 'remedial' or 'organizational' - and is focused on the individual. The group provides a means by which an individual's problems can be assessed and addressed. It is most clearly connected with social work and casework and case management. The emphasis is upon 'ameliorating or preventing the adverse conditions that negatively influence individuals and result in deviant behaviour' (Reid 1981: p191). Classic

examples of this literature come from Gisela Konopka (Konopka 1963) and Paul Glasser et al. (Glasser, Sarri, & Vinter 1974).

Interaction –Focused Groupwork - here the group is understood as 'a system of mutual aid wherein the worker and the members are engaged on the common enterprise of carrying out the group's goals' (Reid 1981: p191). Within this category fall humanistic approaches such as those of Glassman and Kates (Glassman & Kates 1990), the social groupwork of Grace Coyle (Coyle 1937) and the work of William Schwartz (Schwartz & Zalba 1971). Part of the problem within this area is that much of the literature remains wedded to a 'helping', rather than 'learning', orientation. This latter orientation links to a tradition of practice that can be described as 'developmental'.

Social Skills Groupwork - where the focus is on dealing with 'those problems that are related to the social order and the social value orientation in small groups' (Reid 1981: 202). This long established set of traditions of practice is closely linked to community organization/community work. Examples from the literature include (Butler & Wintram 1991) and the classic text of Goetschius (Goetschius 1969).

Educational Groupwork – Major approaches that relate to group work in the classroom are known by different labels: cooperative learning, student team learning, group investigation, and collaborative learning. While each of these approaches may differ in certain aspects of learning and instructional design, such as group structure and teacher role, there are three general types of group work: informal learning groups, formal learning groups, and study teams (adapted from Johnson, Johnson, and Smith (Johnson, Johnson, & Smith 1991).

A quick glance at these categories highlights a number of issues. First, and foremost, they orient us to difference. The obvious danger here is that we then overlook something of great significance - the considerable overlap and sharing of dispositions, ideas and practices across the approaches. For example, some elements of AGI content could be categorized as concerned with 'social group work' (with an emphasis on social functioning and interaction) rather than therapy. Equally, the CSG intervention although different in aim and content, by necessity (the groups were voluntary) will include elements of group boundaries and sensitivity training more usually associated with group therapy.

Second, the categories are not fully discrete, and some systematic approaches to groupwork can belong in different categories according to how we 'cut the cake'. Take

someone like Carl Rogers for example (Rogers 1970). At one dimension he falls within a therapeutic framework, at another his humanistic emphasis leads us to interactionalism. Lawrence Shulman, by focusing on 'helping' could be seen as case-oriented, but his focus on mutual aid places him with an interactionalist approach (Schulman 1979).

There is a problem with categorisation that recurs in the description of the groupwork investigated as part of this study. A large part of this problem results from the manner in which the interventions described below have evolved in an 'adapt to fit' way to co-exist with the school system.

Common Features of CSG and AGI interventions

In a social psychological sense, a group 'is a plurality of persons who interact with each other in a given context more than they interact with anyone else.' (Sprott 1958 p9). Thus the criterion of relatively exclusive interaction in a given context is the principle feature of psychological groups. There is an implicit recognition of this description in the features common to both the AGI and CSG. Each intervention ran for an hour a week for 12 weeks. Each 12-week intervention programme involved 8 children. The same practitioner conducted both group sessions and he was equally qualified to do so being both a teacher and a therapist. The AGI and CSG groups shared many features in the way they were constituted and were designed to provide the context for the establishment of positive group identity, group cohesion, shared group goals and supportive norms for behaviour (Brown 2000). This was achieved through the adoption of a framework for behaviour and clearly articulated aims that depended on group performance over individual attainment. In this way the dynamics of the group emphasised the co-operative over the competitive in a way that has increasing recognition in realising group goals (Stewart, Manz, & Sims 1999), (West 1996). The manner of leadership for the groups was one of facilitation rather than instruction, with individuals encouraged to help each other in progressing through group content with successful group members rewarded for assistance given to weaker group members (Levi 2001). The emphasis in both groups was on positive enforcement of desired skills and behaviours and on adult warmth and attention for children. A positive group identity was fostered in each group and group achievements were celebrated in feedback to classteachers as well as individual attainment. The experience was structured to foster positive social skills and to be a positive experience for those concerned (Rogerson

1996). The progress of both groups was monitored through marking of produced work (CSG) and management reports reflecting on the perceptions of the facilitator (AGI). Weekly appointments were made with the classteacher in order to deliver feedback about the process and progress of the sessions. Roughly equal time was given in these feedback sessions to children in CSG and AGI groups. During the feedback the needs of individual participants were discussed, as were strategies for containing behaviours inside and outside the classroom. Other individuals were mentioned because of their mildly concerning behaviours and the transition between class and group was a central concern. Although primarily verbal in nature, the classteacher was allowed access to the management reports if desired and was given a copy of the summary report produced at the end of the 12 weeks of intervention. The feedback sessions were optional, but taken up by the majority of teachers on most weeks, developing content in line with achievement in the group (CSG) and highlighting the academic and social difficulties amongst children as part of professional standards in undertaking a groups of this nature.

Differences between AGI and CSG interventions

As well as having important common features it is important to stress the differences between the interventions. These differences lie in the nature of activity undertaken by each group, and the extent to which the therapeutic elements of closed groupwork was encouraged through group process. These differences are the subject of the next two sections where the detail of the interventions is laid out. The differences are summarised in the reflective commentary section.

Section 3.1: The Curriculum Study Group (CSG)

The CSG condition was designed as an intervention to run alongside the Action GroupSkills Intervention in order to provide some control measures for the therapeutic elements in the AGI condition. This was done with the aim of demonstrating that the active therapeutic elements provide enhanced effects in terms of socialisation and adjustment. It is important that the CSG group is regarded as an active intervention in itself, rather than a no-intervention control, or even a second-rate intervention control. Not only is the ethical drive to do the best for the CSG group felt strongly, but also through the provision of a small group situation, there is a philosophical and pedagogical attempt made to maximise the adaptive processes present in the group (such as the

opportunities for peer learning and positive adult attention). In the CSG group as in the AGI, there was a clear intention of developing the abilities of the group, and impacting on their challenging behaviours.

In this respect, the theoretical basis of the CSG intervention can be found in the tradition of Educational Groupwork outlined above. This tradition is based around the insight that students learn best when they are actively involved in the process of learning.

Researchers report that, regardless of the subject matter, students working in small groups tend to learn more of what is taught and retain it longer than when the same content is presented in other instructional formats. Evidence that students who work in collaborative groups also appear more satisfied with their classes is not hard to find: e.g. (Slavin 1980) and especially relating to higher education (Chickering & Gamson 1991), (Beckman 1990).

Of the three types of educational groupwork identified from Johnson, Johnson and Smith (1991 p79), the CSG intervention is a good example of a 'Study Team' described as follows: "*Study teams* are long-term groups (usually existing over the course of a semester) with stable membership whose primary responsibility is to provide members with support, encouragement, and assistance in completing course requirements and assignments. Study teams also inform their members about lectures and assignments when someone has missed a session. The larger the class and the more complex the subject matter, the more valuable study teams can be." (p79ff)

This weight of published opinion has not translated into a commensurate amount of practice in the classroom. What seems to have happened in practice is that teachers have taken on board Plowden's views on having children work in groups, but have preferred to retain individualisation rather than co-operation in that context. See e.g. (Bennett, Rolheiser, & Stevahn 1991)

It is difficult to see how the current educational climate could give the lie to this characterisation, and this is what makes the CSG intervention a valuable intervention in its own right as well as providing a measure of control to the AGI intervention. Setting up 'study team' groups in the schools of North Tyneside was part of the challenge of delivering a successful research process and done with an eye to the

application of collaborative groupwork principles to the classroom (Christie et al. 2001) and (Lyle 1996). The CSG groups consisted of 8 children involved in curriculum tasks running in parallel to the classes from which they were drawn. These group sessions were distinguished from other withdrawal activities occurring in the life of a school by consistency in applying the guidelines governing their function. These consisted of:

Planning for each stage of group work. When putting together the syllabus with the classteacher, topics, themes, or projects were chosen to lend themselves to formal group work. Thought was given about how to organize students into groups, help groups negotiate among themselves, provide feedback to the groups, and evaluate the products of group work.

Carefully explaining to the class how the groups will operate and how work would be marked. The overall as well as the week-to-week objectives of the group tasks were explained and relevant concepts defined. In addition to a well-defined task, every group needs a way of getting started, a way of knowing when its task is done, and some guidance about the participation of members. It is important that group members realise that although withdrawn from class the work they do contributes to school merit systems.

Part of the group aim is to give the children the skills they need to succeed in groups. Those who have never worked in collaborative learning groups and may need practice in such skills as active and tolerant listening, helping one another in mastering content, giving and receiving constructive criticism, and managing disagreements. A point was made of discussing these skills with group members and consistent efforts made to model and reinforce them during sessions.

Written contracts were used. These contracts list members' obligations to their group and expectations for tasks completion. These contracts are similar to those used in the AGI groups but they concentrate on common task goals and working to complete curriculum tasks.

Creating the content that matched the group needs as well as being integration with class curricula meant that the actual content varied greatly from school to school from group to group, depending on age and ability. Principles applied across schools and groups to guide the content of group sessions included:

The group tasks were created to require interdependence. The students in a group must perceive that they "sink or swim" together, that each member is responsible to and dependent on all the others, and that one cannot succeed unless all in the group succeed. Knowing that peers are relying on you is a powerful motivator for group work (Kohn, 1986). Strategies for promoting interdependence include specifying common rewards for the group, encouraging students to divide up the task, and formulating tasks that compel students to reach a consensus. (Source: Johnson, Johnson, and Smith, 1991)

The group work was relevant. Students must perceive the group tasks as integral to the class curriculum, not just busywork.

Tasks were created to for children's skills and abilities. Early in the term, relatively easy tasks were introduced, which were quickly differentiated as members engaged at their own level of cognitive ability. In addition as members became become more knowledgeable, there was an increase the difficulty level. It is important to reward achievement in the same differentiated way, however.

Wherever possible, group tasks allowed for a fair division of labour.

Attempts were made to structure the tasks so that each group member could make an equal contribution. For example, one group is asked to contribute to a project on alternative energy sources. Each member of the group is responsible for research on one source, and then all the members work together to incorporate the individual contributions into the final report. Another group is asked to prepare a "medieval newspaper." Students research aspects of life in the Middle Ages, and each student contributes one major article for the newspaper, which includes news stories, feature stories, and editorials. Students conduct their research independently and use group meetings to share information, edit articles, proofread, and design the pages.

This last example gives a flavour of a couple of the tasks set for different groups. In general it is possible to say that the content of the CSG sessions was curriculum focussed with activities set by the classteacher with the academic needs of the group in mind. There was no desire to characterise this group as having educational special needs, and indeed to do so was not possible as the groups contained very able pupils who found the solitary achievement based activity of the classroom very conducive. This having been said, the content of the sessions was delivered in a manner very close to that of a SEN group, adhering to conventions of theory and practice commonly used in the support group (Clark, Dyson, & Millward 1998). Although the content differed from school to school, and was directed to the explicit development of academic skills, the intention was the utilisation of the demonstrated link between academic achievement and self-concept to impact on disaffected behaviours through improving self-esteem (see e.g. D'Amico & Cardaci 2003). Common themes in the content were maintained; the first being a short project that lasted over three or four weeks giving continuity to the sessions and an identity to the group. Children were also engaged in individual work that was differentiated for ability, and the completed work was incorporated back into the classroom through displays or inclusion in personal folders. One of the short projects was introduced to the class as a group presentation on a static display.

Throughout the sessions the group was treated with high respect, members were encouraged in helping explaining things to their peers and positive feedback was used wherever possible, in a manner that was not de-motivating (Henderlong & Lepper 2002). In this way the group worked on national curriculum Maths and English in an additional curriculum developed by the school classteacher. These sessions were conducted in an atmosphere of emotional warmth, which was used as part of the stimulus-response framework to encourage desired behaviours and encourage a positive social conditioning in the group members. The sessions were developmental in delivery, involving a progression through problems and exercises characterised by an interesting presentation and an applied content. Much was made of the opportunities for peer tutoring made possible by the mixture of ability and temperament present in the group, and groups were conducted with humour and a fully supportive approach from the facilitator. This involved also more individual tutoring than is possible in mainstream classrooms but also whole group teaching. The progress of the group was monitored

through marking of produced work and through weekly feedback to the classteacher about developing content in line with achievement in the group. A reflective diary was not kept with this intervention, partly for practical reasons that the data generated, added to that of a similar nature generated in delivery of the AGI intervention would be too unwieldy, but partly to control for the therapeutic processes that the study is seeking to limit to the AGI intervention. In other words there is a danger of both interventions becoming too similar if the CSG group makes any explicit attempt to enter into a therapeutic framework.

Section 3.2: The Action GroupSkills Intervention

By providing children with the opportunity to reflect on experience, and through legitimising the process through sustained attention, the AGI is designed to allow children to feel more at ease with their internal and external states. This familiarity should dispose them to be more open to the expectations of the socialising environment. The experience of children outside the classroom may be affecting all aspects of school performance, but school rarely provides the sufficient opportunity to deal with negative experiences in the home or on the street, or to celebrate significant events in the child's life. The intention of the AGI sessions was to make a systematic and thoughtful programme where children could develop the skills of self-expression and then to use them to reflect on experience. In this sense the AGI sessions are thought of as constituting therapeutic groupwork. As with the CSG group, attention was spent on the setting up process. Although both interventions could be said to be broadly therapeutic in nature, extra care was taken in setting up basic group conditions for the AGI group as this was seen as important in establishing the conditions on which therapeutic work could grow.

3.2.1: Procedures for Setting up a Therapeutic Group

The structure a groupworker establishes for a group is recognised to have a profound effect on the interactions within the group and its general dynamics (see e.g. Dwivedi, Lawton & Hogan 1993). Therefore, the first element in establishing effective and successful groupwork conforming to the broad AGI aims above concerns the planning and preparation that needs to take place before the group starts. In order to initiate a group and activate the therapeutic potential within a group the AGI places a priority on

working with the school establishment in order to give the group every chance of success once it starts. These basic conditions are factors impacting on the group's life that can be controlled at the beginning of the group's life: Selection, Space and Procedure. Each of these will involve the school in a certain amount of education about the group aims and therefore cannot simply be regarded as group factors.

Selection is the first of these conditions and the most important. Selecting the children with a view to the success of the group itself is one of the key predictors of satisfactory outcome. When the selection of children is skewed by pressures from within the school (e.g. relief of the classteacher or need to be seen to make provision for a particular pupil group) the workability of the group becomes difficult and the positive outcomes reduce. This topic was covered in the previous section, but is here reiterated to underline the impact it has on the every aspect of the group activity. Selection of children for the therapeutic AGI sessions was made according to the criteria laid out in section 3.2.1 to include children with identified need in relation to behaviours that hindered, or had the potential to hinder, their progress in school and life. In parallel to those considerations strenuous efforts were made to manage the membership of AGI and CSG groups so that volatile temperaments and clashing personalities were kept apart where possible. Likewise, bullies and their victims were kept apart as were members of close cliques and exclusive in-groups. Pupils were moved within bands between the AGI and CSG groups to adhere to the principles of randomisation. As discussed below, a group needs to bond in order to activate the normative processes of group function and this was seen as a precondition of effective therapeutic groupwork.

The space made available for the group to meet is important. Space is often at a premium in school and the quality, accessibility and exclusive use of the room will depend on the priority placed on the running of the group by teachers within the school. For running AGI sessions core comfort conditions are important if therapeutic work is to take place. It can be helpful to imagine a hierarchy of needs such as that proposed by Maslow (1968). According to such analyses, the basic conditions of warmth, safety, and bodily needs will need attention if there is to be any progress in the so-called higher needs of positive self- and peer-regard. Experience does give some sense to this: standing outside the room with a bunch of rowdy children while waiting for a key, or having a procession of staff and children rooting about for classroom resources as the

session is underway does nothing for establishing a therapeutic space. Therefore the accommodation needs for the AGI sessions were presented as being a spacious room, well heated, close to the normal class environment, with available keys and exclusive use for the duration of the weekly sessions.

Procedure of setting up in the school is the final precondition of successful therapeutic work and could be classed as the management of stakeholders in the programme. In this category are all the conversations to be had with headteachers, classteachers, governors, pupils and parents that have to be satisfactorily completed prior to therapeutic work (and research) to take place. The consent procedure and ethical reflections are presented in Section 3.4, but if it were possible to put a little flesh on the bones of the strictly informative paragraphs included there, it would be to emphasise the conversational and persuasive nature of getting stakeholders on board. The group facilitator had the advantage of being known to a number of schools prior to engaging in this research project, in fact all schools had had experience of the AGI style groupwork delivered before this project was active. Knowing the facilitator, being comfortable with the working methods, and having seen positive outcomes with cohorts of children is a strength of the design in delivering the methodology (time-consuming for staff, intrusive for parents, and a leap of faith for schools).

The importance of these has been emphasised here, and for this reason Hypothesis 9 is given to the examination of the effects of these 'within-school' conditions on the behavioural outcomes of children. Using the AGI dataset it is possible to examine whether perceived helpfulness in achieving these core conditions correlates with behavioural gains (measured T2-T4) attributed to the AGI and CSG programmes. The outcome of such analysis is reported in Chapter Five.

3.2.2: Clear Aims, Boundaries and Behaviour Outcomes

Moving on from the baseline conditions for the setting up of the group, it is possible to begin to describe the ideas that underlie the practice of AGI intervention. The first of these ideas that shape the practice of running therapeutic groups for disaffected youth is the importance of clear aims and boundaries:

The AGI subscribes to the general therapeutic aims of generating an emotional space within which therapeutic factors such as those identified by Yalom (see Yalom 1975) can become active, establishing a culture of playful expression and respect as the mode through which the group operates, and finding positive outcomes to negative patterns of emotional and behavioural response. In order to achieve these aims boundaries for group activity have to be set.

Some of the formal boundaries are set when considering group factors in the set-up process: such as duration, time available, and group membership. Other boundaries need to be established with the group as it forms: these will include the importance of individual commitment, and the agreement in contract form of the behaviours necessary for emotional safety.

General therapeutic aims are supplemented with aims specific to the group, these are designed to guide the group activity and shape its outcome. Solution-focussed approaches to working with children and young people have grown in acceptance and in the interpretation of what constitutes a solution-focus. These aims are developed through conversation with the school, the classteacher and depend on the experience of the facilitator(s) in making success criteria realistic. Group members also identify personal and group aims during group sessions through activities designed to encourage personal reflection.

This mixture of general aims and those specific to the group provide the overarching structure for the group. These aims will be drawn down into session-by-session aims to guide the introduction of activities and the establishment of the therapeutic space. An example of this is given below.

**Table 3.2: Grid of planned week-by-week aims.
(All sessions are of one-hour duration)**

<p>Week 1.</p> <p>Introduction to the group Identification of purpose & skills Playful culture Boundaries and contract</p>	<p>Week 2.</p> <p>Re-contracting Re-examination of purpose Playful culture Creative-expressive</p>	<p>Week 3.</p> <p>Reflection on purpose Personal responses Individual stories Relating purpose to the individual</p>
<p>Week 4.</p> <p>Sharing personal stories Making stories real in group space Possibility of reforming stories Playful culture</p>	<p>Week 5.</p> <p>Mini-closure Skills of dramatic enactment Celebration of process Reflection</p>	<p>Week 6.</p> <p>Half-term break</p>
<p>Week 7.</p> <p>Boundaries and contract Reinforcement of purpose Re-introduction to culture Attention to completed work</p>	<p>Week 8.</p> <p>Dramatic skills Basic enactment of stories Relating enactment to desired skills</p>	<p>Week 9.</p> <p>Elaboration of enacted stories Alternative endings Relating to desired skills Begin closure process</p>
<p>Week 10.</p> <p>Acting 'as-if' skills in place Training for transition Continue closure process Reflection</p>	<p>Week 11.</p> <p>Celebration of completed work Closure</p>	

(Selected lesson plans are included in Section 3.4)

Informing these aims is a positive skills-based understanding of healthy child function as opposed to a notion of disaffection as a set of dysfunctional behaviours to be cut out of a child’s personality. These aims outlined above are informed by a theoretical understanding of disaffection that permeates all aspects of group function, the promotion of positive behaviours. In order to find a complete set of behaviour outcomes to promote through intervention, the task is first to characterise disaffection in a way that is complete. This clarity can found through the analysis of the DSM-IV offered in Section 1.3 which ended by saying...” The emphasis on understanding the actual experiences and behaviours as an alternative model has many attractions for those concerned with behaviour in school. Firstly, using a social aetiology in preference to (but not to the exclusion of) a medical aetiology empowers key social actors to frame change. Secondly, the notion that the diagnostic entities for treatment are not reliable means that

behaviours need not be framed exclusively by diagnostic category, or by a deficit model of dysfunction. In other words it opens the door for the development of a positively framed structure for desired behaviours in youth disaffection.”

Preserving the distinctive and revolutionary nature of the analysis rehearsed above is a challenge when generating desired behaviours. It is all too easy to slip into clichéd talk of respect and empathy as long as respect and empathy mean quiet, biddable children, or to leave unacknowledged major social factors invested in the outcome of behaviour and revert to being another instrument of institutional control. In order to avoid these pitfalls and to preserve the distinct nature of not being driven by the DSM mindset, I have found it useful to think about the skills necessary for full healthy human function, referring to this when developing behaviour outcomes.

Work in this area of the specific skills of positive function is hard to come by but two helpful models are offered here. The first is the pioneering work of Strayhorn (Strayhorn 1988), who developed a competence-based theory that he defines by saying “...the job of those who would promote mental health is 1) to decide what the target persons need to learn to do better, and 2) to pick the ways to influence them towards such learning. “ (p.3). Strayhorn facilitates this process through a system that asks practitioners to choose and enact the most useful cells from a two dimensional matrix formed by a set of competences (or skills) and a set of competence promoting methods (or methods of influence). Clustering skills that relate to a psychodynamic analysis of healthy development, Strayhorn arrives at nine groups of skills on an ascending order of complexity or “abstraction” (p28-29). Developed through more recent publications (Strayhorn 2002a;Strayhorn 2002b), and given an explicit application to the curriculum environment through a recent research programme run by the RSA (Bayliss 2003). Using these sources the following structure is applied to the skills that the Action GroupSkills Intervention seeks to develop in its participants.

A Structure for Skills

Co-operation and kindness

The skills of forming and maintaining social bonds – skills of building relationships, of receiving and giving love.

Handling joint decisions and Interpersonal conflict.

The skills involved in dealing successfully with forces that may weaken or threaten social bonds: separation and interpersonal conflict.

Dealing with Frustration and Taking Pleasure from the Positive

These skills cast the psychologically healthy child as dealing well with frustration, failure and personal misdeeds on one hand and, as taking pleasure in success, favourable events and prosocial deeds on the other.

Facility of Expression and Vocabulary of the Inner World

These skills are those concerned with developing an emotional vocabulary (or emotional intelligence) for inner experience and a repertoire of safe expressive outlets (Salovey & Mayer 1990)

Sustaining Attention and moving towards Autonomy

Focuses on the skills of letting thoughts influence action in an organised way. and the skills involved in working (delaying gratification) and playing (consuming gratification), without the need to involve others.

Seeing Solutions

This group of skills focuses on the use of guiding principles and meanings for one's life, being able to imagine preferred futures

Viewing behaviours as a form of communication can make the process of changing destructive patterns more manageable. This communication by the young person can usefully be seen as a specific gap in a coping skill. Each challenging behaviour can be interpreted as a message about what the young person finds difficult to deal with, and by inference, what skill they need to develop. Put in these commonsense terms this is hardly cutting edge, but consider for a moment the implications of acknowledging that there is a component of behaviour that has to do with communication and a specific development need. This is an important shift away from looking at undesirable behaviours as aberrations that must be excised before socialisation can take place towards a position that seeks at all times to communicate to young people that negotiating emotionally charged social interaction depends on strategies that can be taught. The AGI seeks to add to the emotional resource the young people keep, teaching strategies and skills in order that they may be learnt: practiced in the group sessions 'as-if' they were already part of the repertoire.

3.2.3: Activities to support Aims and Outcomes

In order to support these ideas a carefully constructed menu of activities was offered to the groups, in order to establish the boundaries for behaviour and to rehearse the skills identified as conducive to healthy function. The activities were introduced in an order from more basic skills (e.g. movement/dance and expressive drawing/painting) at the beginning of the programme to more complex skills (e.g. role-play and symbolic drawing/painting) towards the end. This model is one adopted in playtherapy models e.g. Jennings (1980) and is actively used in practice (Ablon 1996; Hansen, Meissler, & Ovens 2000; Johnson et al. 1999; Lewis, Powers, & Newcomer 2002). The evidence-base for effective groupwork models has been covered in some detail in Chapter One. The broadly progressive nature of successful group activity in AGI sessions has evolved from the descriptions of practice above, offering seven headings into which activities could fall.

Containment and respect

These activities allow the facilitator to establish group boundaries and the culture of group activity. These activities will include simple games of a competitive and co-operative variety as well as use of the circle to establish equal rights and turn-taking. These activities allow the facilitator to evaluate the communication behind the behaviours encountered and the identification of skills targeted by future sessions.

Modelling emotions

These activities introduce a basic vocabulary of emotions and explore their physical and verbal expression free from any narrative context.

Culture of play and respect

These activities aim to rediscover a playful and enjoyably pointless capacity within members.

Creative expressive range

These activities aim to enhance the range of creative expression open to children through different media, praising authenticity over sophistication. Media will always include the gathering of stories from life or fantasy from the children

Dramatic play and characterisation

These activities encourage a playful approach to role and scenario through games free from narrative.

Role-play and ritual

These activities introduce a narrative element from the generated stories, moulding them into scenarios, which can be replayed until a more satisfactory structure can be found

Reflection

These activities are introduced to allow a space for the group to evaluate their experience in the group and their impressions of the activities engaged.

The progression through the seven categories is a progression of complexity of task and of the skills required in order to undertake the activity with any success. Therefore the progression is also one of anxiety: the more complex the activity, the greater the chances of failure and the higher the anxiety felt by group members when undertaking the activity. The skilful practitioner will match the group's function with activities providing stimulation but not likely to overwhelm with anxiety.

The following grid uses the activities from the lesson plans at the end of this chapter and for the sake of illustration places them in the structure above. Descriptions of the activities themselves are to be found in the explanatory notes accompanying each lesson plan.

Table 3.2.3a; Structure for Activities

Containment & Respect	Modelling Emotions	Culture of Play and Respect	Creative Expressive Range
<i>Quality Audience</i> <i>Talking Forum</i>	<i>Feelings & Skills</i> <i>Sheets</i> <i>Categorical Groupings</i>	<i>Group Juggle</i> <i>Zip Bop</i> <i>Group Soup</i> <i>Ding a Dong</i>	<i>Near to Far From</i> <i>Magic Newspaper</i> <i>Six-Part Story</i>
Dramatic Play and Characterisation	Role-Play and Ritual	Reflection	
<i>Fruit Salad</i> <i>Pass the Mime</i> <i>Liar Tag</i> <i>Building a Shed</i> <i>Scenario Sculpts</i>	<i>HA!</i> <i>Conflict Dyads</i> <i>Wink Murder</i> <i>De-Roling</i>	<i>Talking Forum</i>	

Activities often fit into more than one group, depending on the introduction and manner in which they are conducted. In addition activities given above (for example Six Part Story) are developmental in themselves, generating material for several sessions that will encompass several categories as activity products are written, drawn, shared and enacted before being reflected on and improved for performance.

The actual activities themselves are considered as not intrinsically of any worth, certainly not as magic potions able to make socialised children out of the disaffected. The AGI approach lays no claim to any comprehensive menu: these are best sought out from the 'how –to' section of the literature, which contains some excellent ideas for group activity (see Mosley & Sonnet 2002; Mosley & Sonnet 2003). The activities in AGI session are regarded as vehicles for the group aims and the relationships essential to the establishment of therapeutic group norms. Equally as important, and linked to the establishment of relationships in the group, is the way the activity allows rehearsal of life skills identified as useful to the group's function.

The evaluation of the worth of any activity is not therefore, 'does it keep the group occupied for 15 minutes, but does it help in the achievement of the group aims? As discussed above, the group aims encompass the therapeutic goals of role repertoire, socialisation etc. and use a positive interpretation of behaviours to identify the skills lacking in the children that make up the group. Obviously there will be some activities that are more effective than others in encouraging the development of skills, but these will vary from group to group.

The methods and application of dramatic approaches in AGI sessions adhered to conventions developed by practitioners wanting to use dramatic enactment to give physical representation to intra-psychic states (Moreno 1964) and extended to include the more simple utilisation of the immediacy of drama in therapeutic settings (Jennings 1987; Emunah 1994). These latter approaches draw out the narrative of a situation close to the experience of individuals in the group, encouraging the use of dramatic techniques to either heighten the experience or to contain it. In this way the insights from narrative therapy are used and enhanced. Sessions would typically begin with a 'talking circle', an opportunity for each child to speak out and be heard. The earlier sessions were characterised by co-operative and competitive games, so that talk would focus on

common themes or experiences within the group. These also served the purpose of creating a group identity, mutual trust and the setting of boundaries for behaviour. The middle sessions used a range of self-expression and controlled physical activity, as well as artistic work such as painting and making of masks. Later sessions utilised the material from this expressive phase to create dramas, enacted by the group for the group. The nature of the drama varied from group to group depending on their experience and the way each group approached expressive materials. For instance, a group might have developed a fantasy play drama using predominantly their mask work or real-life role-plays with children taking on the characters of figures from their world. In the final phase the sessions were characterised by reflection on the learning points offered by the group experience and integration of individual insights into ideas for behaviour outside the group. The children generally determined the wide range of themes that emerged in the groups. For instance, a child might seek to re-enact the sequence of events surrounding a limb fracture but without the frame in which his frantic mother shouted at him. Children might confide that they had been tearing the roof covering off the school and were chased by the police. The group might then rehearse an imaginary sequence in which a transgressor has a dialogue with a 'police officer' and then takes on the 'police officer' role, considering and re-enacting often frightening or disturbing events from different perspectives. Other types of themes that the children chose to explore or come to terms with included disciplinary encounters with teachers, parental separations, and moves of home or school.

In order to illustrate the systematic nature with which the activities are introduced to facilitate the group process, the following table plots the way in which certain activities develop the skills named above. This should not be taken as prescriptive in any form, for so much in terms of outcome depends on the relationship the practitioner has been able to form with the group. Successful activities serve to develop this relationship as well as to engage the group participants.

Table 3.2.3b: Activity Skill Matrix

	Co-operation	Handling Conflict	Pleasure Positive	Facility Expression	Attention Autonomy	Seeing Solutions
<i>Quality Audience</i>	Yes	Yes	No	No	No	No
<i>Talk Forum</i>	Yes	Yes	Yes	Yes	No	No
<i>Feeling & Skills</i>	No	No	Yes	Yes	No	Yes
<i>Categorical groupings</i>	Yes	Yes	Yes	No	No	No
<i>Group Juggle</i>	Yes	Yes	Yes	No	No	No
<i>Zip Bop</i>	Yes	Yes	Yes	No	No	No
<i>Group Soup</i>	Yes	No	Yes	No	Yes	No
<i>Ding a Dong</i>	Yes	No	Yes	No	Yes	No
<i>Near to Far From</i>	Yes	Yes	Yes	No	No	No
<i>Magic Newspaper</i>	Yes	No	Yes	Yes	No	No
<i>Six-Part story</i>	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fruit Salad</i>	Yes	Yes	Yes	Yes	No	No
<i>Pass the Mime</i>	Yes	No	Yes	Yes	Yes	No
<i>Liar Tag</i>	Yes	No	Yes	Yes	Yes	No
<i>Build a Shed</i>	Yes	Yes	Yes	Yes	Yes	No
<i>Scenario Sculpts</i>	Yes	Yes	No	Yes	Yes	Yes
<i>HAI</i>	Yes	No	Yes	No	No	No
<i>Conflict Dyads</i>	Yes	Yes	No	Yes	Yes	Yes
<i>Wink Murder</i>	Yes	Yes	Yes	No	No	No
<i>De-rolling</i>	Yes	No	No	Yes	No	No

The twenty activities for this analysis have been selected as representative of the range that a group might engage with. The list is certainly not exhaustive and in placing them in the table above a lot of the meaning and nuance that a skilled practitioner will use to hold a group's attention is lost. But it is true to say that each activity used in the AGI sessions is subject to some analysis of this kind, and those used to further to the aims of the group are applied systematically as the group advances in its process.

The activities were all introduced with the purpose of encouraging reflection on important relationships and the situations in which relationships are played out, this being an essential element in the transformation of positive group process to therapeutic group process. As reflection was engaged so positive relationships in the group were encouraged through adherence to group rules and in feedback as to the way group members supported each other outside the group. In this way relationship enrichment and resolution of familiar conflict situations were clear aims of the sessions. These elements to the AGI sessions were designed to transfer insights gained in the group to outside situations, and can be seen as an explicit training for transfer. Therapeutic records of each session were kept as a matter of course. These records aided in the planning and targeting of activities in the sessions to meet individual needs. These

records, coded to preserve confidentiality, are designed to aid the facilitator to conduct the internal conversations about children's needs necessary to be effective in the delivery of a therapeutic intervention. In addition it provided a running record of the session progress, a professional measure designed to provide a source of information if the content were ever to be questioned. The group facilitator received supervision for clinical issues from an occupational therapist working in a child and mental health unit

3.2.4: Therapeutic Models

The Action GroupSkills Intervention can be seen as part of a tradition of therapeutic groupwork with young people without necessarily subscribing to this or that therapeutic approach exclusively. The AGI intervention is designed to operate within an education setting, and has a strong focus on mutuality and learning. The therapeutic literature is dominated by material from psychoanalytical or psychodynamic perspectives with a focus on casework. As a consequence, interventions such as the AGI with its emphasis on co-operative learning and problem-solving in the school setting have to borrow from a number of therapeutic models in order to be able to fully articulate the framework through which it can be understood to be a helping group.

Therapeutic groupwork with young people is thought to have been first attempted by Moreno in the early part of the 20th century using psychodrama techniques. In 1934 the Jewish Board of Guardians started groupwork with latency age children, which later became known as Activity Group Therapy. The main aims were to a) improve their ego strength and sense of self worth so often crushed in problem children, b) provide substitute love if they were unable to find this in their homes c) offer opportunities for genuine interest in leisure time activities and d) rebuild their distorted personalities. During the 1960s Ginott (1961) popularised therapeutic groupwork, and the insights of these pioneering practitioners have allowed many techniques to be focussed on the needs of young people.

With much emphasis being placed on the explicit acquisition of life skills including appropriate emotional expression, it is not surprising that the advances in solution-focussed therapies should be of great interest to the running of the AGI. Applying solution focused approaches to groupwork has had recent and persuasive advocates

e.g. (Sharpy 2001), and it is a natural partner to the AGI model that looks to a skills-based approach to overcoming difficulty. The brief therapy model of this approach developed by de Shazer (1985) is a good example of the kind of thinking that has become useful to the running of AGI sessions. Problems and behaviours associated with them are often regarded as static situations in which 'the same thing always happens': the same thing appears then to take on a quality of a rule of life. What happens in the group's life when the problem is not happening can therefore be called not the exception that proves the rule, but the exception to the rule. It is these exceptions that can then be used to construct solution behaviours: the exceptions themselves forming the basis of the solution so that group members may only do more of what they are already doing in order to solve the problem (George, Iveson and Ratner 1999).

These therapies both) regard complexity to be built up over time have a sequential (but not necessarily linear characterisation of the complexity of group activity emphasising non-verbal methods more likely to succeed with young people. The literature in dramatherapy borrows from that of playtherapy when it sets out its conceptual sources as

'...most obviously theatre, and aside from theatre is dramatherapy's immediate predecessor, psychodrama. A third source more primary than theatre in terms of human development is dramatic play, a fourth source, more primary than theatre and psychodrama in terms of the development of civilisation is dramatic ritual. The fifth source, role-play, takes a central position: the experimentation with and embodiment of roles is at the core of dramatherapy' (Emunah 1994 p3).

In AGI sessions therapeutic objectives are formulated in conjunction with an ongoing examination of the particular needs, issues of individual groups. Nonetheless there are certain general goals that arch over particular cases, and still borrowing from the strengths of therapies that do not depend on verbal expression. In AGI sessions these could be articulated to include the following. The first is expression and containment of emotion: feared emotions can be expressed with a sense of safety and distance using dramatic techniques. A related goal will be the development of the reflective capacity: dramatherapists call this the observing self (Deikman 1982), and characterise it as detached from the ego self with the ability to think rationally in the face of emotional

turmoil. A third goal is the expansion of the role repertoire: using drama: real-life limitations on interaction with others are suspended and permission is given to experiment with identity and to practice new ways of relating, leading to an expansion of the self-image. The final goal is the facilitation of social interaction: the collective and collaborative nature of non-verbal expression make drama a good method of practising positive interaction with the hope that the relationships and trust of the group will act as a microcosm for what is possible in the world.

The historical application of therapeutic group factors was generally characterised by a long-term, open-ended, open-content format. However, more recent appraisal of the potential of group approaches has seen a move away from this characterisation (Dies 2000; MacLennan 2000). Together with a recognition of applying solution focused approaches to groupwork (Sharry 2001), and mixing group factors with other treatment approaches (Pollock & Kymissis 2001). Using group theory in this way to frame group content is a statement of intent about the purpose of gathering individuals together, and can be seen framing different treatment modalities (Hayward et al. 2000; Canham & Emanuel 2000; Toren et al. 2000; Snyder, Kymissis, & Kessler 1999).

Section 3.3: Reflective Commentary

This chapter has focussed on the development of the interventions used in this study in the context of the underlying theory. The point made in the last chapter is recalled: “Rotheram-Borus found that although supposedly all effective interventions are theory-based, often the theories are not as comprehensive as the programmes...” (see p40 above). In doing so, a bias towards one intervention does appear, in that the development is presented as more comprehensive for the AGI. However, this apparent bias was subject to corrective measures in the practice of application. As will be made clear in the next chapter, both interventions were equally subject to scrutiny by the school, and equal amounts of feedback were expected on group progress by the classteacher. Indeed the structure of the consent process would not allow one intervention to be felt by child participants and their carers to be inferior in outcome. The processes engaged by the interventions were indeed different, with the reflective capacity encouraged by the AGI and internal states of those involved speculated upon in a way that was not appropriate to the CSG outcomes. This chapter should be regarded

as a guided tour through some of the highlights of group practice and how these relate to both interventions, but how they are magnified through the work of the AGI.

I recall an AGI session in STC School with a group of high(ish) functioning Y6 children unable to completely socialise. The work had taken the group down the path of making masks which they were exploring by holding up to their faces while looking in the mirror. During the course of the session children were invited to sit and give improvised answers to questions as if they were in the character of the mask. One girl, a small impish child with a wicked tongue, held up the mask of an alien and began answering in a voice completely altered from her usual tenor, and giving answers so far removed from the ordinary, that there was an obvious sense amongst the group of entering a different therapeutic space. The group began asking very interesting questions about the life of this alien, and the alien replied using vocabulary that was at once fantastical and ordinary, as she described the feelings of loneliness that resulted from being off-putting to other people. As I watched carefully from the side, I could see this child speaking from a different place, and as I drew the interview to a close, and watched her come back into her normal voice, I hoped that at some level she had learned something about who she was.

It is episodes like the one related above that require an explanation of what processes might be acting on AGI group members to activate reflective capacities. This type of moment was not within the remit of the CSG group, engaged as they were in curriculum studies. This is not to decry the level and effect of their engagement, but rather to highlight the different nature of the interventions and why the challenge of explaining children's experiences of insight is a challenge that required the bulk of the preceding chapter. The experience of running the interventions gives some confidence to the statement that both groups were engaged with vigour and the CSG children often reacted well to the increased attention levels and rewards of praise for real achievements. But as with issues of choice in methodology articulated below, the elements emphasised in the account of theory given above are those with which the author feels most drawn to. There is an inescapable bias due to the fact that it is one person delivering the interventions and explaining their active ingredients and not another.

Section 3.4: Selected Lesson Plans

These lesson plans have been chosen to give an elaboration of how the activities (see previous section) fit in with the aims, and how the aims and activities illustrate the development of the group participants. This development has been given some theoretical structure in the sections above, but broadly speaking, it is hoped that individuals entering the group will be encouraged to put to one side personal attention-seeking strategies for the sake of completing group tasks. Group participants will be encouraged to reveal some simple personal information in a way that is controlled by them. Both to discover that they are not alone in facing their problems, but also in order to gain some insight into the patterns of behaviour that may be adversely affecting them. In order to achieve this, group participants are led through a series of activities that are not the same for every group, but are variations on a theme. These activities engage the group at the level of their tolerance of anxiety, as self-disclosure and new tasks provoke feelings of vulnerability and inadequacy. The development of self-awareness and self understanding are at the end of the arc that the grid above and the lesson plans below describe.

3.4.1: Lesson Plan for Week One

School:

Date:

Attendees:

Aims: Relax and Introduce Group (laughter)
 Begin Feelings and Skills Lists
 Contract and Boundaries
 Movement and Games

Preparation: Chairs in circle
 Area marked by tape on floor
 Juggle balls
 Flipchart and marker pens

Introduction:

“Why Are We Here?” conversation: to reduce anxiety and to set out the aims of the programme.

Talking Forum: introduce self (name) and 1 piece of information about self (see notes)

Game 1: Group Juggle (see notes below)

Game 2: Zip Bop (see notes below)

Development:

Feelings and Skills Sheets (blank) explained and put on wall. (see notes)

Quality Audience (see notes)

Contract – to be signed next week. Includes: Date time frequency, broad aims: to get on better inside and outside school & to work on key life skills, behaviour boundaries: no act or words of violence, confidentiality: what is said in this room stays here (qualified in line with ethical practice).

Work:

Physical boundaries are introduced and the expectations about moving in the space now and in the future. This is done through the activities Group Soup and Near To Far From within the taped area (see notes below).

What follows is designed to reduce anxiety: Ding a Dong (see notes)

Closure:

Talking Forum: How do you feel at this moment?/ one thing you enjoyed from the session.

Pass the Ball Behind the Back (see notes)

HAI (see notes)

Thoughts For Next Week:

Next week will use positive memories of this session to reinforce the boundaries, special, behavioural and expected commitment. Activities will focus on group formation through Individual sharing

Explanatory Notes:

These activities are easy to do but difficult to write down. They all work in practice (I promise), and can be done with all ages of participant. A lot will depend on the ability of the facilitator to communicate a sense of fun and enjoyment when introducing the game – giving permission not to take oneself too seriously.

Quality Audience: This is done as a brainstorming: the group is asked for examples of an audience, theatre, cinema, and classroom. They are then asked what makes an audience a good audience. “What would you see and hear from a quality audience?”

The facilitator is looking for qualities that include specific behaviours and actions on body language and action. When written up this list of behaviours provides a reference for the desired attitudes to be displayed by participants.

Talking Forum: Ritual for start and finish of sessions. Encourages controlled self-disclosure and reflection. Sitting in a circle, the facilitator uses a stone to pass around the participants. When the participant holds the stone he is entitled to a quality audience from the rest of the group. The facilitator introduces the topic of each forum before the stone starts its journey around the group. Topics typically are “Tell me something good that happened to you this week outside school” “Tell me something that you enjoyed and something you would change about this week’s session.” The facilitator may draw out certain contributions through respectful curiosity.

Group Juggle: One ball is passed around the seated group, touching everyone once and not going to anyone twice before being thrown back to facilitator. The point is made about the manner of throwing. It is a measure of ‘groupness’: how well the assembled individuals function as a group, subsuming individual urges to act out to the group aim of completing the task. The ball is then passed around the circle again in the same order and on the third round, an additional ball is introduced. Most groups can manage 2 balls few groups can manage 4.

Zip-Bop: Seated in a circle the group is asked to imagine an electric current passing through the group passed on by a 90-degree movement of the arm towards the person seated next door and the word “Zip”. This current is passed around the group once. When the Zip returns to the facilitator, he introduces a second element whereby the current can be bounced back in the opposite direction by putting two hands up as a shield and saying “Bop”. Inevitably the current gets stuck over one side of the circle, at which time the third element is introduced whereby the current goes across the circle to a recipient clearly indicated by an outstretched arm of the sender along with the word

“Zap”. When zapped, the current may be passed to either side through a Zip or passed across the circle through a Zap. It is important that a Zap cannot be Bopped.

Feelings and Skills Sheets: these sheets provide a resource for group participants.

Whenever a feeling is mentioned as part of the group forum, it is written up on the sheets that can then be referred to in subsequent activities. The skills sheets functions as the same, but also act to make explicit and specific the skills the group aims to target. These will vary from group to group but are always positive and not clichéd.

Group Soup: using a taped off area of the floor, it is explained that this is the space that the participants have to act in during their time in the group. In order to preserve safety (it is explained) we need to be clear about how we move in the space. So children are invited to move around to all corners of the space independently, but without touching others. A close eye is needed and this simple rule repeated when participants bump into each other. They are then asked to greet other participants they pass by nodding their heads and catching their eyes. Finally they are asked to imagine that the floor represents a huge canvas and they are painters: it is their job to create a Jackson Pollock masterpiece by walking their paintbrushes to all corners of the space, moving in every unexpected direction but still keeping to the central rule of not touching.

Near To Far From: following on directly from Group Soup; participants are asked to choose someone from the group but not to name them. It is then the task to move to be as far from that person as is possible in the space while obeying the central rule of not touching others. Emphasise the need to change positions several times. When some kind of equilibrium is reached the facilitator shouts freeze and selected participants asked whom they were trying to move away from. Participants are then asked to choose a different person from the group (still not saying who their choice is), and to move to be as close to that person as possible while obeying the central rule of not touching. When equilibrium is reached, ask one or two of the participants who their choices were. Finally, the participants are asked to move to be as far away from their first choice while at the same time moving to be as close to their second choice. A certain kind of Brownian motion should result after a time the facilitator closes the activity.

Ding a Dong: Sitting in a circle an object is passed from the facilitator to the next person as he passes it, he says, “This is a Ding”. The recipient says, “It’s a what?” The facilitator repeats “It’s a Ding” The first recipient passes the Ding to the next person and says, “This is a Ding” the second recipient says “It’s a what?” At this point, the first recipient ‘forgets’ and says to the facilitator “It’s a what?” The facilitator repeats, “It’s a Ding”. The

second recipient passes the Ding onto the next person in the circle by saying "This is a Ding" the third recipient says "It's a what?" The second recipient 'forgets' and says to the first recipient "It's a what?" the first recipient 'forgets' and says to the facilitator "It's a what?" and the facilitator says, "It's a Ding" This information is passed on to the third recipient. This process repeats until the 10th participant passes the Ding back to the facilitator and the question "It's a what?" passing all the way back around the circle and the answer "It's a Ding" passing all the way back. In the second stage a Ding is passed one way around the circle and a "Dong" passed around the other way. When they cross over the game usually breaks down but not before confusion and laughter.

Pass the Ball Behind the Back: The group stands or sits in a circle with arms clasped behind the back, within arm's reach of each other. One participant stands in the middle and closes her eyes while a small ball is passed behind the backs of the other participants. The participant in the centre opens the eyes and has three guesses to find where the ball is. The ball has to be passed from person to person while she is guessing. Participants in the circle may help each other by pretending to pass the ball.

HAI: is a ritualised ending to each group session. The participants stand in a circle and the facilitator makes a karate-like move stamping his foot and chopping his arm down in the direction of the middle of the circle while loudly shouting HAI After recovering from the shock, the participants are invited to do the same thing but it has to be at precisely the same time as the facilitator repeats the action and sound. The facilitator explains that after this action the group is finished for this week and he looks forward to seeing the participants next week at the same time. He then waits for a moment of concentration amongst the participants before chopping his arm and shouting HAI The participants will do it at the same time and will go out of the door laughing.

3.4.2: Lesson Plan for Week Five

School:

Date:

Attendees:

Aims:

- Prepare group for half term closure (increased anxiety)
- Controlled self-disclosure to find central stories told by group
- Build skills of dramatic enactment
- Controlled, safe closure

Preparation: Feelings and skills sheets on display
Contract and Quality Audience on display
Rolled up newspaper

Introduction:

Talking Forum: "Tell me about one time when you were in trouble." "Did anything help you get out of trouble?"

Categorical Groupings: the aim of this exercise is to encourage self-disclosure in a game format, and foster the therapeutic factor of universality – the feeling that one is not alone in one's troubles. (see notes)

Fruit Salad: this exercise uses the same format as Categorical Groupings, but is used to release tension and to stimulate laughter (see notes).

Development:

Pass the mime: used to develop basic dramatic enactment skills (see notes)

Magic newspaper; used to emphasise the role of the imagination in future group activity (see notes)

Liar Tag; used to combine the skills introduced above (see notes).

Work:

Building a Shed: used to introduce the notion of a performing space in the group and to reduce anxiety about entering it. (see notes)

Scenario sculpts: used to make scenarios out of the stories gathered in the Talking Forum in the form of frozen moments (see notes)

Closure:

Talking Forum: "What have you liked about being in this group?"

Wink Murder: used to provide a safe enjoyable closure activity (see notes).

HAI

Thoughts For Next Week:

Expect behaviours echoing the formation phase in Week One.

Rehearse boundaries

Activities to be low-anxiety

Explanatory Notes:

Categorical groupings: the group sits in a circle and the facilitator explains that a category such as “All those wearing black socks” will be spoken, and those individuals it applies to have to swap seats, with everyone finishing in a different seat to that they started from. The facilitator continues suggesting categories at the level of anxiety that the group is comfortable with. In this version of the game no one is left in the middle and as the group gets used to the game, they are invited to suggest categories themselves. The children remain in control in as far as they do not have to respond to the category suggested even if it applies to them, and the game is aimed at establishing common links between participants as children realise that their situation is mirrored in the rest of the group. Facilitator suggestions might include “All those who have a pet”, “All those who argue with their parents”, and “All those whose parents have split up”.

Fruit Salad: this classic of youth games serves to lighten up the categorical groupings activity if required. The facilitator stands in the middle of the group and pointing to each of the participants in turn, gives them the name of a fruit “Apple, orange, pear. Apple, orange, pear.” At the end of the naming, there should be roughly equal numbers of apples oranges and pears. The facilitator then asks all the apples to change places, then various combinations of two fruits, then says “Fruit salad”, when all the participants change places. To make it more interesting, the children are then given three ways of moving between the chairs: the first being to mime a swimming motion while walking between seats, the second to mime a kangaroo while moving between seats, the third being to mime a pirate while moving between seats (with obligatory wooden leg, eye patch, and parrot on the shoulder). Then apples are invited to move seats as if they were swimming, and oranges as if they were a kangaroo and so on. As the game reaches a close other combinations of multiple fruit and multiple movement are suggested, with a final suggestion of fruit salad moving like a swimming pirate kangaroo. This activity works with all ages, even groups of teachers, and serves to break down barriers in an atmosphere of humour and not taking oneself too seriously.

Pass the Mime: the group sits in a circle and the facilitator explains that in order to engage in the more dramatic of group activities, the participants will have to build up basic acting skills, and these are first and foremost the ability to make others believe that things that are not present, are present. To illustrate this the facilitator mimes a sugar

lump between thumb and forefinger holding it out to the group. "This sugar lump will be passed around the group and when it gets back to me, it will have become the size of a very large beach ball. This growth will be gradual, increasing a little from person to person, and the sugar lump does not get any heavier." This having been accomplished the sugar lump reduces to original size and the facilitator passes it round again with the instruction that as well as getting bigger this sugar lump is getting heavier so that when it arrives back to the starting point, it is the size and weight of a large boulder. This exercise serves to establish turn taking and group co-operation, with the obvious dramatic skills.

Magic Newspaper: building on the pass the mime activity, this activity is introduced as extending the power of the imagination to invent objects and to convince others that they exist. At this point the facilitator takes out a rolled-up newspaper and says something like: "You think this will be difficult, but if I do this...(holds newspaper up to eye and squints down it) you will immediately be able to tell what it is – shout out when you know (repeats action)." The participants will usually have no difficulty in identifying a telescope. The newspaper is then passed around the group with each individual invited to mime an object and the others invited to shout out when they think they have identified the mime. Participants having difficulty getting started can be encouraged to think of things they use every day to get ready for school, something to do with cooking or some sporting thing. Participants who mime a spliff or a penis can be successfully defused by the facilitator saying: "Well there is always one in every group and congratulations this time it is you". The group should be allowed two rounds of the mime.

Liar Tag: The mime theme is continued in this activity, with the level of complexity being quickly built up. The facilitator begins a mime, something simple like cleaning the teeth. The first participant to his left is instructed to ask him in a puzzled way "What are you doing?" the facilitator then continues to mime brushing his teeth but over the course of a few seconds changes the action into another mime similar in action to the first, such as combing his hair. At the same time he says "I'm combing my hair". At this point the first participant begins to mime combing her hair. The second participant to the left asks in a puzzled way "What are you doing?" the first participant changes the mime into a similar mime such as waving goodbye to someone on a train and says "I'm waving goodbye to someone on a train", at which point the second participant starts to mime waving goodbye. In this way the mime makes its way around the circle changing as it does so.

The facilitator's job is to keep the mime changing into a similar action, as the effect of changing it too fast is to fragment the flow of the activity.

Building a Shed: the final stage in these pre-drama activities is to establish the use of a performance space. This is achieved by arranging the chairs in a horseshoe shape and the facilitator standing in front of the group saying "In this space we are going to collaborate to build a shed. I am going to start by measuring out some wood." At this point the facilitator mimes taking a pencil from behind his ear holding up a tape measure and kneeling to mark off a length of wood. The participants are invited to say what else needs to happen in order for a shed to be built. Then with these ideas fresh in their minds the participants are invited to come into the performing space one by one in order starting from the left and saying to those in the space (it will be only the facilitator to begin with) "What are you doing?" Those in the space, and there will be more of them as the activity goes on, reply "We're building a shed" the participant then says "Can I help?" Those in the space immediately reply in unison "Of course you can", the participant then joins the mime by beginning an action that will be useful to the completion of the shed. When everyone is involved, the facilitator encourages everyone to avoid having their back to the horseshoe and steps out of the mime to oversee the action, encouraging a harmonious whole, before saying freeze and miming taking a photo.

Scenario Sculpt: this activity uses the dramatic skills encouraged by previous activities to illuminate actual scenarios that have been shared in the Talking Forum by group participants. Three or four scenarios can be selected and the facilitator will be thinking about which are suitable as the session is begun. At the point when this activity is beginning, the facilitator will recall the scenarios from the Talking forum and selecting a more confident participant, will ask for her permission to look at the scenario she outlined. The participant is invited to an empty chair to the side of the performing space and rehearses the scenario in conversation with the facilitator. During this conversation the facilitator is looking to identify the key players in the drama and the central dramatic moment that in some way encapsulates the scenario. The participant is then asked to choose a person to be her in the sculpt and others to be the remaining three central characters. These are invited into the performing space and sculpted by the 'director' who is aided to get a frozen moment as realistic as possible. Each of the actors is given an attitude and body shape as well as name and relation to the 'director'. Although the director is played by a participant, the scenario is tweaked until it represents something of the reality of the situation. The whole process will take between 5 and 10 minutes

depending on how co-operative and self-conscious the actors are. The sculpt finishes with the facilitator asking the director if there is anything that needs to be changed, asking the actors to freeze and taking an imaginary photo.

Wink Murder: it is a mystery why this game is so popular with groups, it has little intrinsic therapeutic value, but is a great game to close the session because it focuses the attention and gives most participants a chance to be a central player. To begin the game one participant leaves the room and the remaining participants choose a murderer. The first time it is played, it is advisable that the facilitator be the murderer. The participant outside the room is invited in and given the instruction to catch the murderer in three guesses before she kills the other participants by winking at them. This detective has to stand in the middle of the group. When one murderer is caught another leaves the room and the process continues. This can take up to 10 minutes at the end of a session.

3.4.3: Lesson Plan for Week Nine

School:

Date:

Attendees:

Aims: Elaboration of enacted stories from Talking Forum
 Use of rehearsed skills to experiment with alternative endings
 Relating to desired skills
 Begin closure process

Preparation:

Prominent display of Contract, Feelings Sheet, Skills Sheet.

Previous completed work on display

Introduction:

Talking Forum: "What did you get up to this week?" and "Tell me about something you wished you had done differently"

Central Narrative Themes: building on work done in previous two sessions, central themes for the group are drawn out

Conflict Dyads: simple exploration of conflict themes (see notes)

Development:

Six-Part Story (see notes): individual work form completed last week distributed.

Participants invited to cut stories into constituent parts and put into six collective piles.

Generating the Drama: group reminded of all work completed and the conversations that have taken place. Drama constructed using template of six-part story through direct questioning of group: who is hero in this story? What is the environment? In this way character and plot are outlined

Work:

Development of Drama: Ascribing characteristics to characters and plot. Possible brief hot-seating. Initial run through of drama.

Relating Action to Skills: The performance criteria of the drama are quickly given second place to the reflection on the drama. This is done through questioning of group and applying them to the drama. What skills are being displayed here? What skills would be useful here?

Alternative endings: The drama is played through a few times with different endings substituted. These are examined for excitement, and satisfaction.

Closure:

De-roling (see notes)

Talking Forum: How are you feeling at the end of the session? Is there anything you would like to continue with next week? Is there anything you would like to change for next week?

Choice offered for an ending activity

Thoughts For Next Week:

If the drama has life in it, the participants will want to continue experimenting with the forms available to them. Additional impetus can be given by bringing in props or musical instruments.

The mode of expression can be changed if things are stuck by making plasticine models of the characters or environment and creating a miniature play. Or if extra distance is required masks can be made and used.

Explanatory Notes:

Conflict Dyads: two participants are selected and placed on chairs facing each other in the performance space. Each participant is given half of a dyad, examples of which are: I want to go/I want you to stay; I want it/you can't have it. Each participant says her half-dyad in turn for about 5 minutes. As the activity progresses the actors are likely to vary the tone and volume of their interactions and feedback gathered at the end of the 5 minutes about how it felt for the actors and audience at various points and how it might be done differently. Three or four dyads are done in this way.

Six Part Story: this is an activity for developing narrative and character. Participants are invited to divide a piece of paper into six roughly equal spaces. In the first they draw the hero/heroine of the story, in the second space they draw the environment this hero inhabits, in the third they represent the task or quest, in the fourth they draw the problem or obstacle the hero has to overcome, in the fifth the help he gets and in the final space the resolution of the story. These storyboards are completed in one session, then shared in the next session and used in the third session.

De-roling: a simple exercise completed after taking on a character in a dramatic setting. In its basic form this is a matter of each actor saying "I am not so-and-so's mother, I am David".

Chapter Four: Methodology

Introduction

This chapter sets out the practical implications of the position set out in the discussions presented in the chapters above. The intention is to give an account of the methodological principles that guide this research study towards reliability and replicability. Much time and energy has been spent designing the methodological protocols for engaging with the subjects and the data produced. These protocols have emerged from an epistemological position that is explained below but is one that has set some high value in rigour and control procedures such as those set out in the discussion of the Cochrane criteria in Chapter One. In addition this chapter seeks to give an account of the instinctive choices that have provided the initial impetus for the research, and the interaction between these sympathies, and the strict methodological position set out below. The aim is to provide some clarity about the checks and balances that redress the instincts and biases to allow methodological rigour. This chapter should contain everything that is essential for the replication of this study, and through being honest about the competing claims of rigour and bias it should be possible to navigate through issues of bias. Any team can only bring its own strengths to the table, and these will be expressed most forcefully in the instinctive primary choices that get a project underway. For example, the response of T. Quibell to the villains of MWP School or to the expletive-laden chain-smoker of JSP School can only come from the training, background and personality at his disposal. That the response was a therapeutic group situation using creative media is a given, as is the inclination towards the experiment as a method of establishing causality. There might have been other responses, but the research would have then used those as its set of givens. These givens are framed by the principles that have allowed the literature to be sorted, presented and critiqued. Principles such as those presented in the Cochrane criteria (see above) have allowed the central questions of the research to be framed, and in the pages that follow steps taken to acknowledge any bias are seen redress bias. This chapter begins by making a case for considering the practical issues as the primary focus in considering the concerns of epistemology, for these basic and lofty considerations are subject to bias located in the realities of getting things done.

Section 4.1: Research Design:

4.1.1: Relating conceptual frameworks to AGI Design

Conceptual frameworks are here taken to mean... “The form our understanding and knowledge takes, the assumptions that underlie it and the methods used to establish it” (Stevens 1996). Using bigger words it is possible to say these are matters of epistemology and ontology, of pushing back into philosophy to attempt some clarity about the influences on the conduct of research. Bryman suggests the following factors as influences on the conduct of social research (Bryman 2001):

- A. Theory
- B. Epistemology
- C. Ontology
- D. Values
- E. Practical Considerations

Consider the implications of progressing through the list: theory precedes epistemology that precedes ontology then values and practical considerations. This is assumed to be the way that research is constructed and therefore it is the way that an account of research should be constructed also: after immersing oneself in epistemology, ontology and comparative methodology, one arrives at a considered and correctly critical judgement as to the best fit for the research study that one wishes to launch into the world.

Is it not closer to reality to say that one already possesses a set of epistemological, ontological and methodological preferences (admittedly not usually couched in those terms) that has been informed by hidden factors such as one’s prejudices, expertise, training, beliefs, experiences, why one wants to do research, why one sees the need for research in the chosen area, who the subjects are, what will the research achieve, who the audience is for the finished product, how much of the research is to inform practice, how is it orientated to empowering the research subjects? Is it not better, more honest, to acknowledge that the epistemology and ontology, the grand theories of how our worlds are put together can be used to justify the position we started out from?

The logic of this argument is that to reflect the shape of the research process, writings on research would reflect the bottom-up order of the influences on the conduct of social research. This would mean that chapters like this would begin with practical considerations before considering the values of the research process then going on to look at ontology, then using epistemology to observe how the framework for knowledge has been determined by the factors that went before. This does not mean to imply that abstract arguments have not been determining factors to the structure of the AGI study, merely that the practical considerations merit more consideration than is usually given to them.

In the AGI study, there is a strong feeling that the philosophy of knowledge cannot be written off by using social reality as the prime determinant of knowledge and value. There is a strong feeling that the considerations of epistemology do have a value independent of their genesis in the bottom-up framework, a genesis achieved as the result of a number of other factors (listed above). But the important point being made here is that these considerations of epistemology are considered to become activated after the shape of research has been determined by practical considerations and values.

E. Practical Considerations

One interesting consequence of inverting the traditional order in which to consider the factors affecting the conduct of social research is that it becomes apparent that not all the factors are treated as equally important. (Bryman 2001) devotes just 2 pages to practical considerations and values from a chapter of 24 pages. The rest are devoted to theory, epistemology and ontology. This is interesting because it clearly exposes a value system at play, a value system and an epistemology. An epistemology of epistemology if one wishes to bend one's mind in that way. It is almost as if, having exhausted the energy on the deep considerations of the knowledge and the knowledge of the knowledge, there is no energy left to consider the practicalities of circumstance in the research process nor the values of the research team. This makes it doubly difficult to refer one's epistemological conclusions to one's research design, as researchers have to do. It makes it difficult to refer the second half to the first half because half of the reference is not structured to the same degree as the other half.

Bryman does not say much about practical considerations, and what he does say is in the most general terms: 'While practical considerations may seem rather mundane and uninteresting compared with the lofty realm inhabited by the philosophical debates surrounding such discussions about epistemology and ontology, they are important ones. All social research is a coming together of the ideal and the feasible. Because of this there will be many circumstances in which the nature of the topic or of the subjects of an investigation and the constraints on a researcher loom large in decisions about how best to proceed.' (Bryman 2001) p24). Those 3 lines cover a multitude of sins and bear poor testimony to the barriers that a researcher must overcome and the issues she must actively negotiate in order to begin and then complete a successful research process.

For the Action GroupSkills Intervention study (AGI) the practical considerations that had a direct affect on the research design can be grouped as those to do with:

1. The Research Team
2. Subjects and data sources
3. Resources
4. Audience

The Research Team: at the stage of research design, the team consisted of Toby Quibell, supervised by Dr Paul McArdle a consultant child psychiatrist and David Moseley, a reader in Applied Psychology at the University of Newcastle. Of these three, it was Toby Quibell (TQ) who was active in delivering the research and the interventions evaluated through the research. In essence it was to provide an academic foundation to his innovative work with disturbed and disaffected young people on the Meadow Well Estate that the research was undertaken. So the research had to focus on young people and the effects of an intervention delivered in the school context. It was in reality limited by the skills-base of TQ in intervening with children, and his ability in managing a data-gathering process. The point here is that the research reported in this study did not start as a discussion amongst academics as to how best to further the knowledge of disturbed children, but in the cut-and-thrust of practical service delivery. This meant that a practical research design suitable for use in the community had to be chosen.

Subjects and data sources: the subject group had already been selected through the setting up a service delivery to children in school. As minors, their participation in a research programme would be determined by their carers' perceptions of the research process and this meant that a non-intrusive design had to be used. In addition because of the status of children and their reliability in reporting their own gains in behaviour and affect, the research had to use multiple sources of data including the school records and parent and teacher reports.

Resources: the availability of money has a determining effect on all research programmes. Where the money comes from, how much it is and how long it is given for all impact on the choice of research design: indeed, they impact so heavily that often the word choice is often a misnomer. Research design becomes what is possible under the circumstances to achieve the research aims. The AGI study was funded initially from the same resources that paid the wage of TQ. The quality of the initial process added weight to an ultimately successful bid to the NHS regional Research and Development Unit that funded the follow-up part of the research design. The pertinent points here are the two-stage funding process that meant that a simple research design could be developed longitudinally, and that the initial paucity of funds could not support grand designs, but any design had to have the possibility of being enlarged. In addition to money as the chief consideration in this section, other resources that have a determining effect of choice of research design are time and energy of key members of the research team. The time spent putting together the NHS bid was considerable and had to be included by PMcA in a busy consultant programme of service delivery. Similarly for TQ, the initial research process had to be incorporated into the service delivery programme. If neither person had the energy for the task, the design would have looked quite different.

Audience: key to this research, and to the design of this programme was the audience that the research outputs had to convince. As mentioned, the research was undertaken to provide evidence able to determine the efficacy of a pre-existing service delivery. Anecdotal evidence from schools, children and parents supported the AGI groupwork programme as a useful intervention for children in school showing disaffected behaviour. As such those involved with the AGI programme were seeking to convince a wider audience of its value in order to increase the geographical remit of the service. The audience would include children, teachers, parents, headteachers, healthcare workers,

doctors, peers in education and therapy, editors of journals, academics, policy-makers in the Department for Education and Skills, the Department of Health and the Home Office. To persuade these audiences the research design would have to provide a range of data from a range of sources, but would include data of a high validity and reliability. In addition to the need for evidence with the ability to persuade (should the outcomes be positive), the research reporting must be persuasive. This last criterion extends to this piece of writing. It is important to be able to write with an idea of what the audience wants to know, giving a fair account of the research while being alive to the pressures of positive reporting. At the most basic level the audience, including the reader of these words, will want to gather information about whether the study is honest, and whether the research has been conducted in keeping with its core values and the reader must be encouraged by the style of writing to persevere to the conclusions.

D. Values

The values referred to here are the values of those making their choices about research design under the practical constraints outlined above. The values reflect either the personal beliefs or the feelings of the researcher. Bryman states that 'On the face of it, we would expect that social scientists should be value free and objective in their research. After all, one might want to argue that research that simply reflected the personal biases of its practitioners could not be considered valid and scientific because it was bound up with the subjectivities of its practitioners.' He goes on to say that '...there is a growing recognition that it is not feasible to keep the values that a researcher holds totally in check. These can intrude at any or all of a number of points in the process of social research: choice of research area; formulation of research question; choice of method; formulation of research design and data collection techniques; implementation of data collection; analysis of data; interpretation of data; conclusions.' (Bryman 2001) p22).

Recently researchers are prepared to forewarn readers of biases and assumptions and how these might have influenced the subsequent findings, in other words how the values held by the research team have intruded into the research. There is a certain inevitability to this, and a feature of recent research has been the efforts made by the research team to 'confess' the preconceptions that might have affected their integration with the research area. Indeed while reflexivity is still seen as a particular concern of qualitative

research there is a strong argument that research on all minority groups (including those diagnosed with medical conditions and those socially excluded), needs to be consistent with the wider political needs of the subject group. For example, (Mies 1993) has argued in feminist research the 'postulate of value-free research, of neutrality and indifference towards the research objects, has to be replaced with conscious partiality, which is achieved through partial identification with the research objects' (p68). Here the notion of the inevitability of researcher values intruding into the research is to be embraced, rather than held at arm's length with a sheepish grin. Work with unglamorous minority groups requires a particular skill set and this will often include strongly held views, which will have a direct effect on the research design.

But it is at this point that the academic discussion becomes muted, almost as if the discussion of values is alive only within the section of the methodology entitled 'values', and 'reflexivity'. It is possible to see part of the limiting factor for this debate being the relative position that such considerations have in the epistemology-ontology hierarchy. It is more erudite to ascribe the deciding factors of research design to the lofty realms of epistemology etc, rather than determined by mundane factors of practicality and traditionally frowned-upon factors of researcher beliefs. Embracing the values held by the research teams involves exploring how those beliefs impact on the design of the research. Finding a reference point for these values is often difficult. For the AGI team, there are few better places to start than the pedagogic creed of John Dewey. Six statements from this article sum up the position from which the AGI study begins:

'I believe that this educational process has two sides, one psychological and one sociological; and that neither can be subordinated to the other or neglected without evil results following.'

'To prepare [the child] for the future life means to give him command of himself; it means so to train him that he will have the full and ready use of all his capacities'

'I believe that the school is primarily a social institution. Education being a social process, the school is simply that form of community life in which all those agencies are concentrated that will be most effective in bringing the child to share in the inherited resources of the race, and to use his own powers for social ends'

'I believe that only through the continual and sympathetic observation of childhood's interests can the adult enter into the child's life and see what it is ready for, and upon what material it could work most readily and fruitfully.

'I believe that if we can only secure right habits of action and thought, with reference to the good, the true, and the beautiful, the emotions will for the most part take case of themselves.'

'I believe that all reforms that rest simply upon the enactment of law, or the threatening of certain penalties, or upon changes in mechanical or outward arrangements, are transitory and futile' (Dewey 1897).

In the AGI study, the practical factors and researcher values had an effect on the research design equal to that of epistemology. Of course it is possible to say that epistemology is implicit in values and beliefs, and that even within practical constraints there is always a choice between competing methods and epistemology has a guiding role in shaping those choices, but here it is important to underline that a researcher's values are produced and embedded over a number of years through processes of chance, experience, upbringing and training and consideration of these complex processes falls somewhat outside the remit of epistemology, but have been given attention in Section 1.7

As the section on ontology and epistemology is engaged, the discussion of research design is a more accurate reflection of the actual process of research design as it happened in the field. There already exist principles under which the epistemology will be ordered. In this section I argued that this is the norm, but other accounts like to present the emergence of principles as happening in the opposite direction. The principle guiding the engagement with questions of ontology and epistemology can be expressed as: the form our understanding takes is determined by our belief that children can change to become more successful in managing their environment and that this change can be measured. Furthermore this change can be communicated to others by means of research methods that are persuasive in establishing that it is not just a chance happening that change has occurred.

C. Ontology

Questions of social ontology are concerned with the nature of social entities. The central point of orientation here is the question of whether social entities can and should be considered objective entities, and have a reality external to social actors, or whether they can and should be considered social constructions built up from the perceptions and actions of social actors. These positions are frequently referred to respectively as objectivism and constructivism. This discussion follows that of Bryman (Bryman 2001) p17), in his illustration of the differences between the two positions by reference to two of the central terms in social science – organisation and culture. Objectivism asserts that social phenomena and their meanings have an existence that is independent of social actors. It implies that social phenomena and the categories that we use in everyday discourse have an existence that is independent or separate from actors. We can discuss an organisation as a tangible object. It has rules and regulations. It adopts standardised procedures for getting things done. People are appointed to different jobs within a division of labour. In thinking in these terms we are tending towards the view that an organisation has a reality independent of those who inhabit it. The organisation is a constraining force that acts on and inhibits its members. The same can be said of culture. They are repositories of widely shared values and customs into which people are socialised so that they can function and full participants. Culture constrains us because we internalise beliefs and values. In the case of both organisation and culture, the social entity in question comes across as something external to the actor and as having an almost tangible reality of its own. It has the characteristics of an object, and so it has an objective reality.

However, we can consider an alternative ontological position, constructivism, which asserts that social phenomena and their meanings are continually being accomplished by social actors. It implies that social phenomena and categories are not only produced through social interaction, but are in a constant state of revision. In recent years the term has come also to include the notion that the researcher's own accounts of the social worlds are constructions. In other words, the researcher always presents a specific version of reality, rather than one that can be regarded as definitive. Knowledge is regarded as indeterminate in a position known as postmodernism. Both the ontological version of constructivism and the post-modern meaning are antithetical to objectivism, and post-modern constructivism is antithetical to realism. Post-modern constructivism

can be thought of usefully as constructivism in relation to the nature of knowledge in the social world, and the ontological meaning as constructivism in relation to the social world – that is one that views social objects as socially constructed.

When examining an organisation from this perspective, there is an alternative to viewing order in organisation as a pre-existing characteristic. Instead it is possible to argue as did Strauss in researching a psychiatric hospital that order is best conceptualised as 'negotiated' and less like commands and more like general understandings (Strauss et al. 1973). Because few areas of the spheres of action were prescribed, the social order of the hospital was an outcome of agreed-upon patterns of action that were themselves the products of negotiations between the different parties involved. Social order is in a constant state of change, and it is argued that a preoccupation with the formal properties of organisation (rules, charts, regulations and roles) tends to neglect the degree to which order in organisations has to be accomplished in everyday interaction. Much of the same kind of point can be made about the idea of culture. Instead of seeing culture as an external reality that acts on and constrains people, it can be taken to be an emergent reality in a continuous state of construction and reconstruction. For example (Becker 1982) has suggested that '...people create culture continuously...No set of cultural understandings...provides a perfectly applicable solution to any problem people have to solve in the course of their day, and therefore they must remake those solutions, adapt their understandings in the light of what is different about it' (p521). Importantly both Becker and Strauss understand that the constructivist position cannot be pushed to the extreme: it is necessary to appreciate that culture has a reality that persists and antedates the participation of particular people, a reality that shapes perspectives, but is not an inert objective reality that only possesses a sense of constraint: it acts as a point of reference but is always in the process of being formed. While neither of the authors discussed propose a constructivist view of organisations nor that of culture that pushes the constructivist argument to the extreme, each admits the pre-existence of culture and organisation, in each case there is a stress on the active role of individuals in the social construction of social reality. This central idea is explored more fully in the work of (Walsh 1972) and (Giddens 1984) for example, and can be seen particularly in discourse analysis. As Potter observes: 'the world...is constituted in one way or another as people talk it, write it and argue it' (Potter 1996 p98).

Questions of social ontology cannot be divorced from issues concerning the conduct of social research. If epistemology can be characterised as the nature of knowledge – how do we know? – then ontology can be characterised as the nature of what is known. Ontological assumptions and commitments will feed into the ways in which research is carried out. If a research problem is formulated in such a way as to suggest that organisations and cultures are objective social entities that act on individuals the researcher is likely to emphasise the formal properties of organisations or the beliefs and values of members of a culture. If, on the other hand, the problem is formulated so that the tenuousness of organisation and culture is stressed, or that groups of behaviour characteristics are constructed into categories, it is likely that an emphasis will be placed on the active involvement of people in reality construction. In either case, it might be supposed that different approaches to the design of research and the collection of data are required. Even so, as we consider the polarities of social ontology outlined above in the light of the principles that already govern the shape of the knowledge acceptable to the AGI study, we find that relating the two discussions is problematic. We have a research programme rooted in the practical and fully acknowledging its value-laden nature. The moral imperative is strong; the call for change that should be heard through all the writings about the research is urgent. How then are we to relate this to a discussion of ontology? To put it another way: how does our ontological position reflect our moral imperative? Again the important point is to work bottom-up through the practicalities and values rather than top down in adopting an ontological stance and fitting research to that. Our central belief is that change is possible in children's behaviour through an instructive intervention based on the enrichment of the skills-set of children. Furthermore this change we believe to be measurable. There is a persuasive case to say that clustering symptoms into diagnostic categories for which there are corresponding treatment procedures is an incomplete conceptualisation of the issues facing children, and the diagnostic categories are constructs, the acceptance of which overlooks their shaky foundations as objective realities. As discussed above, Bentall (2003) gives a convincing account of the strengths and weaknesses of the diagnostic tradition that can be illustrated through the ADHD debate. By gathering symptoms together and giving them the category name ADHD the education and medical community has created a syndrome for which the objective basis in physiology is debateable. ADHD can also be seen as a construct, an aggregation of skills shortages. To view it in the latter way makes the cluster of symptoms more open to change

because the aetiology has more possibilities. In this way our understandings are closer to the constructivist position: that illness is created continuously, no set of understandings provides a perfectly applicable solution to any illness and so solutions must be remade, adapted in the light of what is different about it. In the search for the ontological shape to carry our beliefs, we understand, like Becker and Strauss that the constructivist position can be pushed to the extreme and the position outlined here falls some way short of the logical conclusion that knowledge is nothing more than a product of a particular cultural stance and hence no more valid than witchcraft. Thoughtful criticism of the social constructivist approach (Greenwood 1994) points out that it is important to examine the way that cultural and historical forces have distorted our scientific theories, but argues that we should use the insights gained to improve our knowledge and develop better theories. From this position of critical realism it is possible to accept that school culture and family culture have a reality that predates the participation of particular people, and the same can be said of the symptoms of disaffection amongst children. It is hoped that our design reflects this position with regard to social realities, acknowledging that change is often ephemeral and patterns difficult to discern and attention is better spent attempting to measure behaviours rather than constructs.

B. Epistemology

If ontology can be characterised as the nature of what is known, then epistemology can be characterised as the nature of knowledge – how do we know? Epistemology then, concerns the question of what is or should be acceptable knowledge in a discipline. A central issue in the context of this research is whether the social world can and should be studied according to the same principles, procedures and ethos as the natural sciences. The natural science epistemology has become known as positivism. Central tenets vary between authors but will generally include the following:

1. Only phenomena and hence knowledge confirmed by the senses can genuinely be warranted as knowledge.
2. The purpose of theory is to generate hypotheses that can be tested and that will thereby allow explanations of laws to be assessed (deductivism).
3. Knowledge is arrived at by the gathering of facts that provide the basis of laws (inductivism).

4. Science must (and can) be conducted in a way that is value free (that is, objective).
5. There is a clear distinction between scientific statements and normative statements. The former are the true domain of the scientist. This last principle is implied by the first because the truth or otherwise of normative statements cannot be confirmed by the senses.

Positivism contains elements of deductive strategies (2) and inductive (3), and draws a sharp distinction between theory and research. The role of research is to test theories and provide material for the development of laws. There is the strong implication that it is possible to collect observations in a manner that is not influenced by pre-existing theories. Moreover theoretical terms that are not directly observable are not considered genuinely scientific entities, and there is therefore the implication that greater epistemological status is given to observation rather than theory.

Although it is close to truism to say that there has been a drift away from characterising scientific practice in terms of positivism, there is always the difficulty in determining whether the drift away is away from the scientific approach generally, or the philosophical term of positivism itself. For example realism, more particularly critical realism purports to provide an alternative account of the nature of scientific practice (Bhaskar 1989). While realism shares features with positivism, namely that the natural and social sciences should apply the same kinds of approach to data collection and to explanation, and a commitment to the view that there is an external reality to which scientists direct their attention (that there is a reality separate from our descriptions of it), the specific manifesto of critical realism is to recognise the reality of the natural world and the events and discourses of the social world, holding that “we will only be able to understand and so change the social world if we identify the structures at work that generate those events and discourses...These structures are not immediately apparent in the observable pattern of events; they can only be identified through the practical and theoretical work of the social sciences” (Bhaskar 1989 p2). Critical realism implies firstly that whereas positivists take the view that the scientists’ conceptualisation of reality actually directly reflects that reality, realists argue that the scientists’ conceptualisation is simply a way of knowing that reality. Secondly by implication, critical realists are happy to admit that their explanations contain elements that are not observable. As a result,

hypothetical entities that account for regularities in the natural or social orders (the generative mechanisms the Baskar refers to) are perfectly admissible for realists but not for positivists. 'What makes critical realism critical is that the identification of generative mechanisms offers the prospect of introducing changes that can transform the status quo' (Bryman 2001) p17).

Interpretivism is taken to denote an alternative to the positivist orthodoxy. It is predicated on the view that a strategy is required that respects the differences between people and the objects of the natural sciences and therefore requires the social scientist to grasp the subjective meaning of social action. Von Wright (1971) has depicted the epistemological clash as being between positivism and hermeneutics (a term drawn from theology that is concerned with the theory and method of interpretation of human action). Positivists desire explanation, interpretivists desire understanding. The initial application of phenomenological ideas to the social sciences is attributed to Alfred Schutz (Schutz 1962). Schutz's position asserts that there is a fundamental difference between the subject matter of the natural and the social sciences and that an epistemology is required that will reflect and capitalise on that difference. The fundamental difference resides in the fact that social reality is an arena for the structuring of priority values for human beings and therefore human action is meaningful. Secondly, it is the job of the social scientist to gain access to people's common sense thinking, and hence to interpret their actions and their social world from their point of view. It is this last feature that social scientists claiming allegiance to phenomenology have typically emphasised: 'the phenomenologist views human behaviour...as a product of how people interpret the world...In order to grasp the meaning of a person's behaviour, *the phenomenologist attempts to see things from the person's point of view*, (Bogdan & Taylor 1975).

The hermeneutic-phenomenological tradition does not exhaust the intellectual influences on interpretivism, but what is considered important here is that taking an interpretive stance can mean that the researcher comes up with surprising findings, or at least findings that appear surprising if a largely external stance is taken, an external stance being one outside the particular social context being studied. Also, in adopting an interpretivist stance the researcher is not simply laying bare how members of a social group interpret the world around them. The researcher will be aiming to place the interpretations made into a social scientific frame. There is a double interpretation going on; the researcher is providing an interpretation of others' interpretations. Indeed, there

is a third level of interpretation going on, because the researcher's interpretations have to be interpreted in terms of the concepts, theories and literature of a discipline.

'However, while such interconnections between epistemological issues and research practice exist, it is important not to overstate them, since they represent tendencies rather than definitive points of correspondence. Thus, particular epistemological principles and research practices do not necessarily go hand in hand in a neat unambiguous manner' (Bryman 2001).

This last statement highlights the complexity of pinning down a conceptual framework for a research project such as the one attempted in these pages. Positivism and Interpretivism mingle in a way that is difficult to disentangle, as layers of interpretation lie on top of one another. As the conceptual framework for this study takes shape through the current discussion, this should not be surprising; the epistemology is complex because the subject matter is complex. As discussed elsewhere, children's function and their development is not only complex, but multi-modal and there is no satisfying grand theory that an epistemology can hitch itself to. The chief concern of this piece of applied social research is to forge continuity between the interpretations of subject matter, observations of subjects, and the research methods that allow some kind of reflection on what has preceded and helped in the formation of recommendations for practice. Of paramount importance therefore are not the competing and complementary claims of epistemological approaches, but that the research process is congruent with the frameworks within which children and their development are understood. These frameworks can be characterised as pluralistic and dynamic, accounting for competing claims to truth. In the same way it will be found that the organising system or theory of conceptual framework for this research project will hold together unlikely bedfellows because that bears a better likeness of the complexity of the social and theoretical world inhabited by us all.

4.1.2: Theory in the Bottom-up Conceptual Framework

This approach inverted the traditional order of tackling factors that determine the choice of conceptual framework. These are:

- A. Theory
- B. Epistemology

- C. Ontology
- D. Values
- E. Practical Considerations

In the previous section these factors were considered from the bottom-up. The rationale for doing so was that the order of putting together a research programme was more closely followed by this inverted order. We begin with our idea, look at the practical considerations, make choices within these constraints about design in keeping with our values and then look for knowledge that helps us to establish our case as it emerges. In the course of the discussion in the previous section it was stated plainly that the research was an evaluation of work already taking place, that this work was based on strongly held values about what was a desirable state for children's function, that the research was undertaken in order to provide data upon which the audiences receiving the research could base value judgements about the effectiveness of the work and its suitability for wider development. It was discussed how these factors made demands upon the research design, favouring a longitudinal design and a design able to show generalisability, at the same time as respecting the subject communities and making every effort to bring benefit back to them. Ontologically our position was more towards constructivism than objectivism, and when gathered, data would reflect the fragmented nature of reality, and the contextual nature of behaviours. Epistemologically the position is complex, because social phenomena are complex, but statements about the nature of knowledge should reflect the notion that a pluralistic understanding of the subjects should be reflected in a pluralistic and flexible theory of the nature of knowledge.

Holding the bottom-up factors together in some kind of sense is not easy. In a complex world, looking at complex behaviours and reaching towards explanations that must reflect that complexity, the implication that a notion of choice can apply to a conceptual framework governing a research programme can be viewed as compromised. If account is to be taken of the arguments reviewed above, then the idea that at this point the researcher must come out as a positivist, or a constructivist is unrealistic. As discussed above many of the categories for epistemology actually incorporate competing categories, and while there may be affinities between forms of ontology and epistemology, there is no hard and fast protocol for combining ontology and epistemology. The point here is that far from choosing a conceptual framework, the

framework has to be constructed. In order to do this while being alive to the tautological possibilities in constructing questions to be answered by data fitting the researcher's worldview, organising principles have been proposed, for example by Stevens: Stevens' trimodal theory makes a systematic incorporation of competing methodological techniques possible. And although this system, like any analogy, breaks down at a certain point, it is helpful in giving a structure to the construction of an epistemology that allows the interrelation of factors that otherwise may be regarded as oppositional: it shows for example how behaviour is both biologically and socially constructed, and how people are both determined and yet capable of autonomy (Stevens 1996).

Supported by these discussions, an epistemology for the AGI study can be constructed that is in keeping with the use in Chapter One of integrated theoretical frameworks to interpret children's behaviour. Behaviour is complex, and has a wilful habit of escaping from efforts to encase it in explanation and any theoretical framework for understanding behaviour can be judged on its ability to accommodate that complexity. By the same token, the conceptual framework governing research epistemology that can account most fully for the complexity of behaviour will be best placed to provide data that make a difference.

The AGI study wishes to construct a reasoned argument for using methods of data gathering, analysis and presentation from complementary traditions and frameworks. This is intended to make the research able to impact more widely. Quantitative methods are used, but these are not considered incompatible with a hermeneutic and holistic understanding of children, because the methods are held in a complex epistemology. In any complex system like the one described in relation to children's behaviour, it is impossible to identify causal sequences, to explain why changes come about, but by choosing a large sample and using analytic methods it is possible to make inferences from a global level about what change is happening. This should not be considered as excluding the insights that might be gained from a more obviously reflexive-mode activity such as psychotherapy. Observation protocols in the AGI Study are used to generate another type of data than post-hoc questionnaire reports, data to identify behaviours not constructs, and look at the behaviours generalised between settings. The observation data used a standard procedure and was scored to be represented quantifiably, but the

numerical representation was not intended to obscure entirely the complexity of behaviour itself.

Finally, the research is aimed at providing evidence to change patterns of practice, and therefore must use the power of epistemology and methodology to do so. The data, the evidence and the methods used to gather them are not regarded as having intrinsic value: they are considered to be of value to the extent that people will listen to them. And different people will listen to different types of evidence. The arenas in which change is sought are many, ranging from teachers in classrooms to Whitehall agencies to academic debate in learned journals. For this aim of the research programme to be realised, the evidence will have to be powerful in these different domains, and this means accounting for the different notions of what constitutes powerful research.

So for this study the conceptual framework will be flexible. Where explanations use biological processes to account for behaviours, the research will test hypotheses to look for causal patterns and test predictions. Where explanations focus on the behaviour made as a response to social patterns, the research will test hypotheses to give an account that tries to make sense of what the informants are doing, feeling or thinking, and to relate this account to the one given by the informants themselves through questionnaire. Finally where explanations look for what is possible in behaviour and what are the factors restricting self-awareness and reflexivity the research will seek to make a moral case for promoting change and self-awareness.

4.1.3: The Concerns of AGI Methodology

One of the chief concerns of the AGI study is to deliver data that can be subjected to analyses acceptable to a number of audiences considered essential in promoting structural change. To this end, the AGI methodology has been chosen for its rigour, and analytic work has been done in order that it can be demonstrated that the work has value and the researcher has not been misleading. It is worth reflecting on these principles for a moment to clarify that the AGI study is serious about providing reliable data through considered methods. There are ways in which it is possible to subscribe to the principles above while keeping hidden aspects of the research that are difficult to

explain, providing information that is incomplete and vague, making it difficult for the reader can draw his/her own conclusions.

For example it is possible to follow the well-established methodological and statistical framework for recognising quantitative rigour, without recognising that all methods have their limitations, and that by themselves do not necessarily exclude sloppy research. It is possible to produce a piece of work that qualifies as rigorous but is mendacious, cryptic and vague. These fault lines in the integrity of a piece of research do not have to be intentional: consider for a moment that the randomised controlled trial does not guarantee clinical generalisability, scientific advance or replicability; see (Holmberg, Baum, & Adami 1999; Sohn 1998). There are now many sources of similar rules and conventions that offer comparable rigour in qualitative research as discussed in the previous section (Elliott, Fischer, & Rennie 1999; Mays & Pope 1995; Turpin et al. 1997), and qualitative researchers sometimes cite strenuous methods such as purposive sampling, grounded theory, multiple coding, triangulation and respondent validation. It can be argued (Salmon 2003; Barbour 2001) that when regarded purely as methods, each of these can degrade rather than enrich research, because if engaged without a critical and honest mindset, it is possible to use the products of the method to promote inappropriate conclusions.

Consider also the way analysis is used: we expect to see that a researcher has not merely recorded and meticulously reported data, but that these data have been analysed. It has long been recognised that there is a real danger of having too much of a good thing when it comes to analysis, but the role of analysis particularly in quantitative data has reached a point where the ease of statistical procedures calls into question the value of generating analyses that outweigh the original data itself. In a world where the volume of analysis generates a need for further post-hoc analysis and readers are often left drowning in a sea of correlations and comparisons, less is definitely more.

Paradoxically, it is perhaps to qualitative data analysis that quantitative researchers must look. It is here that researchers have worked at developing an analysis until it 'works' – i.e. exerts the rhetorical weight that is in proportion to the moment of the findings.

Examples include achieving coherence and organisation or empowering the reader or participants (Stiles 1993). Much quantitative analysis would do well to recognise that analytic work is much more than statistical analysis and where necessary it is the quality

of the quantitative over the quantity that must be regarded with primacy. With this in mind the researcher will do well to find ways of organising results that indicate coherence or that focus selectively on findings that will empower readers or participants. It is manifestly true that in the absence of clear theoretical rationale and analytic thought before 'Choose Variables' is selected from the SPSS drop-down menu, computer analysis packages will be no substitute. The question mark is not over analytic work per se but over what work, to what ends and at what time in the presentation process.

Both the methods and analysis can be engaged with procedural propriety but deliver research products that do not take sufficient account of the internal limitations of their own methods and implicit variables of researcher intent that guide the conclusions. It is important therefore that the researcher should not try to mislead. This is a central tenet of quality research. Since Rosenthal demonstrated investigator effects on research outcomes (Rosenthal 1966), the reality that results of psychological research reflect the researcher as much as the researched has been inescapable. In an argument familiar to the qualitative researcher, there is strong evidence for a universal appreciation of the scientist as a participant in, rather than observer of, the field of study (Potter 1996). Although qualitative researchers have tried to reclaim this ground by arguing that the researcher's experience, situation and motivation are integral to their research, there is still the insistence that science is an altogether loftier goal than the biases of politics or rhetoric. Unfortunately, according to Salmon, the practice of accepted methodologies 'whether by the familiar practices that claim to ensure objectivity and reproducibility of quantitative research, or by the emphasis in qualitative research on disclosure of researcher's situation and perspective' (Salmon 2003 p26) are not foolproof in defining the boundaries between science, journalism and rhetoric. Evidence of this is found in those who say that the reproducibility of quantitative findings is not assured by analytic techniques (Sohn 1998) and is rarely tested in practice (Potter 1996) and qualitative researchers' emphasis on reflexivity can merely be a rhetorical device for persuasive effect (Yardley 1997). Therefore, the boundary between science and other kinds of persuasion remains ill defined (Potter 1996), even though we apply it daily. The consumer of the research, the audience, is therefore left to make a moral as much as a technical judgement that the researcher has not been 'too' biased. If the reader's attention is not brought to this realisation, the research stands an increased chance of being incompletely interrogated.

Following on from this it should be possible to know when work is worthless but is this really as possible as it should be? Quantitative researchers are used to presenting their aims as testable hypotheses. In reality they are often not true hypotheses but written after the data have been analysed, and in response to the findings that have emerged. This presents as having dubious intentions, but nevertheless can be seen as part of the research cycle of clarifying aims and exposing the coherence of the finished work to scrutiny, exposing whether it has a message, and allowing the reader to judge the work. Aims that are trivial, unachievable or discordant with the design invite judgement that the work was futile. Qualitative research where it appears should be careful to adhere to the same standards, but all too often qualitative researchers are less helpful in helping the reader in deciding whether the work has any value, taking refuge in ill-defined metaphors to express the aims of the study. For example; the oft stated aim 'to gain a deeper understanding of' defines the arbiter of success as the researcher not the reader, in contrast aims to 'describe', 'show how', and 'develop some concept to understand' might more readily empower the reader to judge whether they had been achieved. Where qualitative researchers have made this process opaque is where the ability to judge the quality of work is in crisis, and post-hoc work presented as hypotheses interrogating data stifles debate about what can legitimately be sought from quantitative data.

This train of thought gives rise to a set of principles to operate alongside those traditionally associated with quality research. These principles include the points raised above and advocate that the work should matter to others not just the researcher. An obvious point, but consider for a moment what proportion of research meets this simple criterion. Research is not, by definition, a self-indulgent activity for the researcher. It has to matter to others. Indeed, if Salmon is right in saying that research quality cannot come from adherence to methodological rules (see above), then it must instead be agreed by an audience – a scientific community (Rennie 2000). This view leads to two further considerations for evaluating research. The first is whether it is clear what the audience can take from it. Quantitative researchers rarely, in the view of Salmon, address the issue explicitly, often sheltering behind the generalisability of their findings. However there is a need for greater honesty about the products of quantitative research. For example, Cronbach, whose alpha coefficient is widely regarded as the stamp of permanence and generalisability on a questionnaire, warned that social phenomena are

too context-bound to permit generalisation (Cronbach 1975). Distinguishing scientific from statistical inference (Holmberg, Baum, & Adami 1999) is not the solution. 'Even if results are used to build and test theory rather than directly to generalise implications to a broader population, the theory's generalisability is at issue' (Salmon 2003 p27). One complication in evaluating the message of quantitative research is that it is almost invariably written up in a theory testing way that is, as noted above, often a pretence. Given the exploratory nature of much statistical analysis, and the 'habitual nature of Type 1 errors' (Salmon 2003 p27), perhaps quantitative research should be judged according to its ability to produce and develop ideas rather than to test them. Quantitative researchers are quick to disown generalisability, but are often less clear about what to put in its place. Accepting that work is 'exploratory' is not a complete answer; it describes the motivation of the researcher, and not what the receiver might take from it. Similarly reference to transferability in qualitative literature ducks the question unless it is clear what is being transferred. There is no shortage of possibilities: concepts developed in a study can equip researchers or practitioners to think or act differently in future; findings can help to identify new hypotheses, or findings can simply be incompatible with prevailing assumptions. What transfers is the need to rethink those assumptions. Whatever it is that is transferred, there is a simpler need for clarity in the body of the research. On this subject it may be worth considering how it is that a practical outlook seems to act as a filter to research; witness how the practical orientation of medicine can be used to explain why qualitative research has been accepted with so little controversy – clinicians can simply judge for themselves when a finding is useful. (Corbin 1998; Macnaughton 1996). Indeed it has been proposed that in medical research synthesis, the message's importance should be weighted, not just the method's rigour (Edwards, Russell, & Stott 1998). This is not to suggest that medical practitioners are in the habit of, or routinely do ignore the (e.g.) Cochrane criteria (Clarke & Oxman 1999), but simply that a pragmatic, practice-based approach to the interface between research and practice, while often unacknowledged can be found to actively complement a consideration of which studies adhere most closely to pre-formed ideas of rigour.

A second consideration arises from acknowledging that research is for others: if this is true, work that never reaches its intended audience – because it is not published or presented – or work that once disseminated is never read or understood or has no

impact on anything or anyone, could not be regarded as research. It is hardly controversial that the views of a scientific community define what counts as research, but it is less readily accepted that the recognition is a social act: the community accepts what it agrees to accept. This raises some interesting issues in evaluating research; for instance, if dissemination to an audience is a component of the research process can student research be said to be complete and examinable before they have demonstrated their ability to communicate it effectively?

This refreshing critique rests on values that emerge from Feyerabend's anarchist perspective (Feyerabend 1975), and articulated well by Salmon, from whom the complementary principles are taken (Salmon 2003). It is possible to object that these values are arbitrary or partial, neither deduced from an epistemological position nor induced from a factor or thematic analysis. But, from an anarchist perspective this is precisely the point. They illustrate an epistemology that is based on subjectivity and behaviour, the test of which is whether truths emerge that reveal themselves as self-evident in the light of pragmatic analysis. In the brief discussion above, the case is felt to be strong for applying this pragmatic 'anarchist perspective' to the sacred criteria of research evaluation.

It is possible to become over-enthusiastic in redressing the balance, however, and bearing all of the above in mind when setting out the concerns of the AGI methodology, it is still important to consider what are the general criteria for admissible data for this research programme. Data gathered will have to be persuasive, and will have to demonstrate reliability, replicability and validity. While recognising that these terms relate better to traditional notions of quantitative data, any qualitative element will be subject to these notions too.

Reliability is concerned with the question of whether the results of a study are repeatable. The term is commonly used in relation to the question of whether the measures that are devised for concepts in the social sciences (such as poverty, behaviour, racial prejudice etc) are consistent. Reliability is a particular concern for those quantitative elements in this research programme, as measures used in questionnaires have to be stable across respondents to have persuasive power. Reliability and replicability go hand in hand, and in order to be properly persuasive it should be possible

to replicate the research findings. This is true even when data will not be gathered in the same way. Teachers in the classroom will want to feel the beneficial effects of behaviour change, even if they do not collect controlled data.

Validity is a wider term that is applicable to research in general; it is concerned with the integrity of the conclusions that are generated from a piece of research. The main types of validity typically distinguished are: *Measurement Validity*, a criterion that applies mostly to quantitative research and to the search for measures of social scientific concepts. Measurement validity is often referred to as construct validity and essentially is to do with the question of whether a measure that is derived from a concept really does reflect the concept that it is supposed to be denoting. Does the IQ test really measure variations in intelligence? Does the Achenbach Internalising scale really measure depressive behaviours? Measurement validity is related to reliability, if an item is unstable, it will not serve as a measurement of the concept it points to. In other words the assessment of measurement validity presupposes that a measure is reliable. *Internal Validity* relates mainly to the issue of causality, the question of whether a conclusion that incorporates a causal relationship between two or more variables holds water. If it is suggested that x causes y, can we be sure that x is responsible for variation in y and not something else that is apparently causing a causal relationship? In discussing issues of causality the factor that has a causal impact is called the independent variable and the effect is called the dependent variable, and internal validity raises the question: how confident can we be that the independent variable really is responsible to the variation that has been identified in the dependent variable? *External Validity* is concerned with the question of whether the results of a study can be generalised beyond the specific research context. It is in this context that the issue of how people were selected to participate in the research becomes crucial, and this is the reason that quantitative researchers are so keen to generate representative samples. *Ecological Validity* is concerned with the question of whether the social scientific findings are applicable to people's everyday natural settings. As Cicourel has put it 'Do our instruments capture the daily life, conditions, opinions, values, attitudes and knowledge base of those we study as expressed in their natural habitat?' (Cicourel 1982). This validity is concerned with the question of whether social research sometimes produces findings that may be technically valid, but have little to do with people's lives. If findings are ecologically invalid, they are artefacts of the social scientist's arsenal of data collection and analytic

tools. For example, the conclusions deriving from a study using questionnaires may have measurement validity and a reasonable level of internal validity and may be externally valid in the sense that the findings can be generalised to other samples confronted by the same questionnaire, but the unnaturalness of the fact of having to answer a questionnaire may mean that the findings have limited ecological validity.

There is no argument here that explicit examination of the criteria by which research is assessed is central to the strength of the discipline (Devers 1999), but simply that methodology needs a debate that is grounded in the untidy and intuitive reality of what we do when we evaluate research, rather in the more ordered and oversimplified logic of what we write to justify what we do. There is a stronger need to be open about the subjective and intuitive criteria that we use to judge all research, and that guides our own research choices. The AGI study makes extensive use of quantitative data and methods of analysis, for reasons presented above and an experimental paradigm was used to guide the methodological choices. These choices were modified over time, in the manner of action research, in order to boost the validity of the study to the broadest audience. The Achenbach questionnaire battery was the first choice of assessment (Achenbach 1993), because of its wide use in the child and adolescent mental health and in published studies. The problem scales of the Achenbach questionnaires do not contain structured responses to positive situations and the Multidimensional Self-Concept Scale (Bracken 1992) was used to supplement the problem behaviour constructs. The Bracken scale also contains a positive measure of family function, and this held out the possibility of triangulation with the parent questionnaire (the Child Behaviour Checklist) from Achenbach. Further additions to the assessment battery were sought to compensate for the reliance of other people's reports of children's behaviour, and so an observation protocol was developed for use in the classroom, providing an independent measure of children's behaviour in a class context.

4.1.4: Choice of Methodology and Justification

After these practical considerations shaped the design, the programme could be described as follows: a community-based longitudinal evaluation of the impact of a brief intervention delivered in schools to a cohort of disaffected children. Data were gathered from multiple sources and compared at time points. The data contain qualitative elements for sensitivity and quantitative measures for reliability.

For pragmatic reasons an experimental design was favoured, but the ethical difficulties in forming a no-intervention control group (see Section 3.4) meant that while the experimental paradigm was used to guide research practice, and as many experimental elements were incorporated as possible, the design should be classified as a quasi-experimental. Quasi-experiments are studies that have certain characteristics of experimental designs but do not fulfil all of the criteria for internal validity (Bryman 2001). In this case, the lack of a placebo control casts a certain amount of doubt on the internal validity because extraneous and confounding variables are not completely accounted for. However the lack of artificiality and strength of ecological validity often make quasi-experiments compelling. Historically, quasi-experimental designs have been prominent in evaluation research, which is the tradition where the AGI sits most comfortably. Evaluation research is concerned with the evaluation of social and organizational programmes or interventions. The essential question typically asked by such studies is 'has the intervention achieved its anticipated goals?' Although there are differences in opinion about how the evaluations of the impact of intervention should be carried out there is general recognition of the importance of an in-depth understanding of the context in which an intervention occurs and the diverse viewpoints of stakeholders (Greene, 1994). Programmes frequently draw on the principles of critical realism (see above) which sees the outcome of intervention as the result of generative mechanisms and the contexts of those mechanisms (Pawson, 1997). A feature of such programmes is the examination of the causal factors that inhibit or promote change when an intervention occurs. These features sit quite comfortably with the imperatives of the AGI research with its emphasis that the products of method must serve the aims of the study. And the aims of the study must include the empowerment of the study's consumers and the moral and technical persuasion of those amongst the audience who act as arbiters of quality. The rest of this section details the active elements of design and how they combine to form a quasi-experimental methodology.

Experimental Elements in the AGI Study

The independent variable is the group treatment approach. The children's performance on a number of scales, completed by different informants is measured: these are the dependent variables. The research was designed so that any change in the dependent variables can be assumed to be the result of changes in the independent variable. The

AGI research uses two levels of independent variable, that is to say, two forms of treatment approach: the Action GroupSkills Intervention and a Curriculum Skills Group. Both levels of independent variable resemble each other in every respect other than the content of the intervention. The only variable that has been manipulated is the groupwork approach. This research project has been designed so that the changes in dependent variables will distinguish between levels of independent variable, allowing conclusions to be drawn about the nature of theoretical understanding of remedial action in antisocial behaviour patterns and the conceptualisation of antisocial behaviour itself.

The data from the changes in dependent variables are predicted to distinguish between levels of the independent variable and the link between the change in dependent and the different levels of independent and the underlying theory will only hold if it can be demonstrated that the presence of confounding variables has been kept to acceptable levels. Extraneous variables are those factors other than the independent variable that affect people's performance on the dependent variables. The logic underlying the experimental method is that changes in the independent variable will cause any change in the dependent variables. This argument depends on ensuring that the only thing that distinguishes between the conditions are the levels of the independent variable. The difficulty with experiments in this area is that seldom can it be certain that it is only the levels in the independent variable that distinguish between conditions. Extraneous variables need to be controlled for and are traditionally divided into three types, participant, investigator and situational. In order to prevent extraneous variables of any category undermining the comparison between levels of the independent variable, four techniques of methodology are applied to the experimental design to allow the argument of causation to proceed.

Emphasis is placed on keeping extraneous variables constant, based on the rationale that if everyone experiences the same influences from the extraneous variable it cannot cause difference in performance under the various experimental conditions. Methods used to achieve this include:

Balancing: this procedure balances variables so that they affect each group equally. Balancing to avoid order effects or practice effects applies to all three categories of extraneous variable.

Randomising: not all variables can be kept constant for all participants. Some respondents may receive the intervention at a different time of day, a different time of year, or they may have differing nature and severity of presenting problem, or it might be necessary to have different investigators collecting data from various groups. The approach to controlling extraneous variables by randomisation is to ensure that unavoidable variations do not help or hinder one group or condition more than another. There will be numerous differences between the people in each group and one way to prevent these having any consistent effect is to allocate people to the groups in a random procedure. The assumption underlying random allocation is that any extraneous participant variables will be more or less equally represented in each group and so will not act as confounding variables. Randomly allocating people to groups is the crucial factor that determines whether the study is following an experimental design.

The Control Group: keeping extraneous variables constant, randomising and balancing are meant to prevent them becoming confounding variables but do not guarantee that extraneous variables have no effect on the respondents. The effects can be measured using an identical group of respondents (a control group) who traditionally receive a dummy intervention. People in the control group are not given the independent variable, so their performance reflects the effects of only the extraneous variables. Comparing the results of the experimental group with those of the control group is traditionally thought to reveal differences due only to the effects of the independent variable. Attempts were made in the AGI study to control for content of intervention, but the ethical principles adopted by the study in keeping with its community focus made it difficult to justify the use of a no-intervention control and crucially this aspect of experimental design is lacking. The AGI study did use a control intervention and adopted a waiting list design where children were assessed 3 months prior to intervention to give the flavour of a no intervention period. This did reveal some unexpected effects and these attempts to provide control data can only be said to have had mixed success.

For these reasons concerning the control group, along with others to do with the community focus and the emphasis on social and policy change, the AGI study cannot be considered a pure experimental design. This does not preclude the incorporation of the following aspects of experimental design to increase the reliability and validity of data collected, as it is considered to improve the possibility of being taken seriously by the desired audience.

A crucial aspect of experimental and quasi-experimental designs is how participants are assigned to the various conditions, how sub-groups are formed. This has yielded four major types of experiment: between subjects, within subjects, matched subjects and mixed designs. Each has advantages and disadvantages.

Between-subjects (randomised groups) Design

In this design people are allocated to form groups and each group is exposed to its own level of the independent variable. The assumptions underlying this design is that the groups are more or less equal before the exposure to the independent variable occurs, so that any differences in the dependent variable can be ascribed to the independent variable and not to any pre-existing differences in the groups. This assumption is accepted if the pool of participants is allocated to the various groups by a random procedure. This does not mean haphazard or uncontrolled allocation, which may well produce biases. In the AGI project randomisation as described in Section 3.2 depended on a random number table selecting children banded according to severity and nature of presenting symptom. Randomising in this way avoids some of the problems associated with matched groups or within-subject designs, but in assuming that subgroups are formed by a random procedure it is assumed that the subgroups will be more or less equal on any measure considered. When however the number of participants is limited (as is the case in most experiments), the groups may still differ; it may be for example, that the sex composition varies. The problems with randomisation in the AGI study are documented, and in cases like these the options are three-fold: first the effects of incomplete randomisation can be ignored and accepted as a factor in the outcome, acknowledging that if the experiment were repeated with a fresh allocation of people to groups then these differences would be unlikely to recur and any possible biases would be as likely to be reversed. Secondly, a within-subjects design may be used in which the same respondents are used in each of the experimental conditions. The third possibility is the use of a matched group design. Neither of these latter two options was viable for the AGI sample.

Conforming to these principles of experimental design by emphasising the elements of methodology considered to deliver validity and reliability gives the AGI study added persuasive power. This is an important consideration for the AGI study, which seeks

change as a result of the research. But the essentially social aspect of this research area means ecological validity is equally important in promoting effective change and the intention at the design phase was to balance the quasi-experimental elements with phenomenological designs to look at the meaning of the intervention experience for those concerned and to complement the persuasive power of the randomised control study with the ecological strengths of qualitative designs.

The action research design with its focus on social change was considered particularly attractive, as the AGI study can be seen as a study of a social situation with a view to improving the quality of the action within. The purpose of the AGI study is to promote change by transforming structures through the influence of the information collected. In action research the question arises out of the problems of practitioners and it is an important aspect of this approach that the analysis of the situation is *in situ*. The immediate aim of the research is to understand these problems, and the researcher, who may or may not be the actual practitioner, formulates speculative and tentative general principles about the problems that have been identified. From these it is possible to generate hypotheses about what action is likely to lead to improvements. However, despite the common ground between AGI and action research models, the phenomenological and qualitative approaches were considered inappropriate on the grounds that the resource implications made them impractical. However, the commitment to these approaches can be discerned in the AGI study through the Classroom Observations Protocols used to gather information about children's behaviour outside the intervention context. This measure cannot be considered qualitative with its emphasis on behavioural categories (talking out of turn, out of seat), but it does use categories of behaviour that have meaning for the children and the teachers, and in this way the influence of a phenomenological inclination may be perceived by those so disposed.

4.1.5: Summary of Methodological Choices

The concerns of the AGI study are to demonstrate that findings can be generalised and to persuade a diverse and traditional audience of the validity of findings. Experimental elements have been used in a quasi-experimental design to generate data that conforms to rigorous performance targets while accommodating the social context of the research and the emphasis on social change. The key features of the design are as follows:

- The study uses two interventions delivered to two randomised groups by the same practitioner. The number of participants is maximised to increase the chances of finding statistical significance.
- The interventions resemble each other in every respect except that of content.
- Data are gathered at time points with longitudinal scope of 1 year
- The core of the dataset is the response made to standardised questionnaires. These responses provide quantitative data for statistical analysis.
- Questionnaire data are supplemented by a direct observation protocol intended to triangulate relevant behaviours with questionnaire data.
- Multiple data sources are used, including the teachers and parents of children as well as the children themselves. The data reflects the children's perceptions of their own behaviour, and perceptions of their behaviour in different contexts. The standardised questionnaires used contain a measure to provide data of a different type to the others: where the Achenbach questionnaires (Achenbach 1991a) provide problem scores, the MSCS (Bracken 1992) gives an account of positive adjustment. In this way the constructed nature of behaviour is acknowledged, if not accounted for. Both these sources are complemented by so-called 'hard' data concerning attendance.
- These different kinds of measure, different respondents and observations are intended to deliver rigour to the design and power to the positive findings.
- A pre-test post-test design is used in which there is a waiting list condition, so that the participants can act as their own controls.

While the AGI has made an ethical attempt to control the intervening variables, it has to be recognised that the intervening variables are actually people, with their emotional responses, their conceptualisations, their needs and their defence mechanisms. The point being made is that the AGI study is an attempt not just to understand the world but also change it. From this standpoint 'objective value-free' research is nonsense, becoming an end in itself and avoiding the real issues of what to do about the problem, what can be done and how to do it.

Section 4.2: Subjects, Instruments and Procedures

4.2.1: Subjects

Description:

The children in this study were all from inner city schools in the North Tyneside local authority that forms part of the Tyne & Wear conurbation. The schools consisted of one Secondary School (Y7 – Y11), one Middle School (Y5 – Y8) and three primary Schools (Reception – Y6). All schools are situated in areas of multiple deprivation and all except the secondary school draw pupils from families characterised by unemployment and low income. The secondary school has a more mixed intake by virtue of its size. The children, when released from class for the AGI sessions were often over-boisterous, in a state of conflict and involved in the kind of politicking that is in continual motion to establish power relationships. They clearly expected the sessions to be a loosening of the reins of school authority and although in the school setting, the children were not minded, nor of the kind of temperament to give the facilitator an easy run. The schools themselves lent a definite flavour to the behaviours of each group. Something to do with the school ethos, perhaps, combined with the conditions and characteristics of the catchment area. For example, the STC primary school is a church school, with a firm ethos of thoughtful discipline. The children took readily to the AGI sessions and undertook complex dramatic activities, but were sulky at one extreme and calculatingly unpleasant at the other. The PSJ children were a younger group, drawn from Y4/5 in a forgotten part of North Shields. The children were quiet, a little nervous, but able to fly off the handle with verbal and physical threats in a way that was impossible to predict. The CEN and MWP schools were in more typically deprived inner-city areas. The children displayed behaviours that can be attributed to multiple factors of deprivation and family breakdown happening together. Curiously in some ways, these children were the easiest to manage in that their behaviours were more or less predictable their limitations were better anticipated, and activities adjusted to increase the chances of positive outcomes. The JSP school was the only secondary school included, and the groups run there had a distinctive character due to the children being a little older: the behaviour was of a type that could be predicted, but happened at a higher volume, with greater potential for damage and unedifying ruckus.

Table 4.2.1a: Subjects

Group	School	Sch.Type P/M/S	No. AGI		No. CSG		Sch. Year	Date of Interv
			M	F	M	F		
1A	STC	P	4	4	5	3	Y6	Spring 1996
1B	CEN	M	4	4	4	4	Y7	Spring 1996
2A	CEN	M	4	4	5	2	Y7	Summer 1996
2B	MWP	P	6	2	3	5	Y6	Summer 1996
2C	JSP	S	4	3	5	4	Y7	Summer 1996
3A	JSP	S	4	2	4	2	Y9	Autumn 1996
3C	PSJ	P	3	3	2	4	Y5/6	Autumn 1996
4A	JSP	S	4	4	4	4	Y7	Spring 1997
4B	MWP	P	6	3	1	7	Y6	Spring 1997
Total n			39	29	33	35		
Total n			68		68			

Table 4.2.1 shows that 61 children were from Y6, 63 from Y7 and 12 from Y9 and that the interventions were delivered over the course of one year starting from January 1996. Both AGI and CSG interventions lasted for 12 weeks, and groups in cohorts 2,3 & 4 were all subject to a waiting list condition where data was gathered one term prior to intervention. In primary and middle schools the proportion of children from one class involved in the group interventions (AGI+CSG) was 50%. This proportion, delivered in schools where one class represents a year group, demonstrates that the group intervention is able to manage the kind of populations included in an at-risk analysis, and is different kind of intervention in the scale of deployment.

Distribution of Initial Scores:

The initial scores for the whole population, AGI and CSG groups were analysed to characterise the location and variability of the dataset and coefficients generated to describe the skewness and kurtosis of YSR, CBCL, TRF and MSCS total t-scores at T1. Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution or dataset is symmetric if it looks the same to the right and left of the centre point. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. That is, datasets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly and have heavy tails. Datasets with low kurtosis tend to have a flat top near the mean rather than sharp peak. A uniform distribution would be the extreme case. A standard normal distribution has a skewness of zero and a kurtosis of zero. The data for the whole population, along with the AGI and CSG groups separately

were found to have skewness and kurtosis values very close to that of a normal distribution, and nothing approaching accepted values for moderate effects for these characterisations (Chambers 1964).

Identification:

The literature of childhood interventions is rich in two areas relevant to this section. The first are outcome studies with children who have been identified by the school family or health systems as deserving of secondary or tertiary care. The second area includes ambitious studies using large numbers in population studies of the incidence of disturbance. There is relatively little in the literature concerning children whose behaviour is of concern to teachers and carers but the severity of which is not enough to warrant the expense of specialist care. This latter group of children contain those who will be consistent users of specialist services in the future, but relatively little by way of research and evidence-based preventative practice occurs with this group. The reasons lie in the methodological and ethical issues involved in identifying and treating children whose behavioural anomalies are not universally acknowledged as worthy of attention by health and education professional and often parents and carers as well.

The clear focus of this study is with this group. The advantages of early intervention are set out in Chapter One, and the acknowledged increase of methodological and ethical issues to be accounted for when compared to specialist populations and epidemiological studies are considered to be a worthwhile investment of energy. Kolvin solved this problem to his own satisfaction by applying a screening battery to 4300 children and selecting 600 for inclusion in the study (Kolvin, Garside, Nicol, Macmillan, Wolstenholme, & Leitch 1981). Junior children (aged 7 and 8 years old or Y3) who 'showed some signs of having social or psychiatric disturbance or learning problems' were considered along with senior schoolchildren (aged 11 and 12 years old or Y7) who definitely showed 'relatively clear-cut psychiatric disturbance' for inclusion in the study. Establishing these cut-offs was a far from straightforward exercise. Kolvin gave space to a lengthy discussion of the variety of problems, circumstances and treatments available, and concludes, "uncovering problems of maladjusted children is a very complex process. In our research we had to develop rapid and effective ways of identifying children in need. We should emphasise that our screen was part of a research enterprise that may not always be best for ordinary practice. Part of putting our research into context will be

discussion of how children should be identified in the normal day-to-day work of the school" (Kolvin 1981 p15). The first point made is that the screening process, although complex, had to make allowances for the rapidity required. This is apparent in the establishment of the cut-off points for including children into the study. The range of instruments used gave Kolvin a bank of contradictory evidence where children would score above the cut-off decided by the team on one measure, but below on another (p336 – 341). Despite the complexity of the screening process, the team was forced into qualitative judgements made on the basis of experience of those involved with the children. The second point made in the quote reproduced above to the effect that the complex screening process is too cumbersome for everyday practice, is an implicit acknowledgement that so-called 'everyday' practice of identification by professionals and carers is an adequate screening tool when conducted thoughtfully and in light of the available evidence from the behaviour of each child.

From the above paragraphs it is to be understood that the precedents for identification in early intervention with limited numbers in a community context are few. The methodological and ethical issues that arise are forbidding and this may account for the paucity of precedent. Hints at how the methodological issues can be solved are given by looking closely at Kolvin, and in fact the current study, limited by resource, relied on teacher screening. The teaching staff of each school was given the list of criteria reproduced in Table 3.2.1b below and asked to make recommendations based on the evidence available to them, formal, informal and intuitive. The teachers concerned always included the head teacher as well as the class teacher and discussions were conducted in triangulation with the research leader.

Table 4.2.1b:Criteria for Inclusion in the GroupSkills Intervention Programme

Group Criteria	The group will be constituted to reflect the range of concerning behaviours in the classroom. All children will satisfy the general criteria and one or more criterion from one or more of the other criteria groups. Care will be taken to balance the group composition when including children with severe symptoms
General Criteria	The group will consist of children whose performance or appearance in school is of concern or who are performing consistently below their potential. This

	behaviour is damaging to their social function.
Conduct Criteria	Recurrent incidence of aggression, intimidation of others, off-task, out of seat, bullying, verbal aggression, inappropriate conduct to teaching staff, truanting, temporary exclusion, unconvincing absence, well-meaning but inappropriate boisterousness
Affective Criteria	Recurrent states of mental absence, sadness, victimisation, deep shyness, low esteem, low confidence, social exclusion by peers
Physical Criteria	Unkempt appearance, lack of care for physical self, overweight, small for age, known major family problems, having physical characteristic that attracts negative peer attention.

The criteria of group composition and severity were the chief concerns in this process. Composing a group of eight children that could be managed is obviously a key to the success of the intervention programme. Experience prior to the current study in the running and planning of creative expressive groups with a therapeutic aim made the communication of these concerns easier. Being in the same room as eight children specially selected for their ability to run about and ignore instructions is a deeply taxing experience, and not one that should be repeated more than a handful of times in one's lifetime. Add to this the need for outcomes to convince outsiders of the efficacy of the programme and a taxing experience become intensely stressful. A properly composed group is the primary condition of successful outcome. A properly composed group should contain a range of symptom and severity, and should reflect the classroom as a microcosm. This is because re-socialisation is the aim and equipping children to survive in a dysfunctional group does not equip them to survive in the 'normative' environment of the classroom and playground (however dysfunctional that may be). The cautionary note about severity does not just cover the difficulties of running a group with severe conduct issues. The nature of the open group process and therapeutic aims of amelioration of internal emotional states does make the discussion of deeply-felt cathartic issues problematic, particularly when the confidentiality of the peer group has to be ensured to minimise further psychological damage. For this reason there is a strong possibility that children with deep social phobia will only experience a recapitulation of their negative experience, and those children unfortunate enough to have a life-experience that has made deep damage to their natural resilience should be considered for provision other than that of a group nature. Quantifying these exclusion criteria is difficult without a screening device: on the other hand there are questions as to whether such behaviour

will be invariably picked up through a screening protocol that was not supplemented by parent and carer information. Also, experienced and sensitive professionals rarely have a problem in identifying the children included in this criterion and because an experienced group worker asks the questions, the lack of a screening instrument for the severity criterion was not a difficulty in running the groups.

Children were not asked to screen themselves, although an opportunity was given to the children to opt out of the interventions prior to their commencement (see below). Asking children to screen themselves more formally is on first sight an attractive option, in keeping with the core values of this study to make children more active decision makers in the processes affecting them. In the balance the arguments against child self-screening were considered more persuasive. Practically the notion is problematic: either involving individual interviews with a school year-group population or the development of a screening tool sensitive enough to identify the children fitting the broad range of criteria above. Added to that difficulty is the subjective judgement that was found to be valuable in constituting the groups that would have to be exercised on the basis of individual interviews of screening scores both of which would be much more numerous. Aside from the methodological issues there is a question as to whether the act of screening children would by its nature exclude some of those fulfilling criteria most clearly. Would those shy, depressed children have the confidence and will to identify themselves as suitable for group intervention, and would children in general have the insight to be reliable in the identification of their internal states? It could be argued that children would be at least as reliable as those involved in the methods described, but it is not the false positives that are a concern, but the false negatives – those unable to see the wood for the trees in their socialising issues, those in a coma-grey of acceptance of an unhealthy social and emotional reality. Finally the presentation of the group interventions outlined above has always generated interest among children and the anticipation would be that there would be many considered suitable for the group, and using a child-completed population screen, there would be many identified as suitable for intervention. The ethical issues in running a limited intervention programme centre around what to do with those identified as suitable but unable to be included.

Allocation to Intervention Condition:

Teachers identified children meeting the selection criteria with numbers limited from each class to sixteen children. There were always enough children identified to make equal numbers in each intervention condition from a single class. In most cases, the number of children identified amounted to 50% of the whole class. The sixteen children were then allocated by the classteacher to one of three categories: conduct, affect or social on the basis of the presenting behaviour of the child. A member of the intervention team, who did not know the children, randomly divided each category into equal numbers of 'A' and 'B' children. 'A' children were allocated to the AGI intervention and 'B' children were allocated to the CSG intervention. After children were allocated, in order to ensure an adequate mix of problems or to avoid a particularly volatile grouping, in consultation with the teacher, children were occasionally exchanged between the groups; the numbers of children affected by a group change was approximately 8 in the whole study. The mean scores of the intervention groups at T1 reveal that the AGI group contained more children with higher scores indicating that post randomisation changes by teachers in group composition favoured the AGI group for children with perceived problems. This approach sought to minimise the risk of bias at the level of the individual, but also, following normal therapeutic group practice, to ensure an appropriate case mix and at the least, the viability of the groups. The age of the children in the curriculum studies groups (mean age 11.3 years, s.d. 1.3) and AGI group (mean age 11.4 years, s.d. 1.1) was similar. In 'tweaking' the groups post randomisation for developmental stage and group viability, the Achenbach scores were not used: the teachers did it on the basis of the experience of the children. Two children refused to participate in the groups for reasons of social reluctance or interest in classroom activities that would have to be forgone to attend group sessions. These children, one from CEN and one from PSJ were replaced with other children from the class meeting the inclusion criteria. Group attendance approached 100% for both interventions

Power considerations were based on the calculations of Cohen (Cohen 1988). In this book, the sample sizes needed to generate various effect sizes of up to one standard deviation are calculated under different conditions of analysis. The power is generically defined as the confidence that can be expressed about being right and by convention 80% is considered to be the correct confidence level. The AGI study is based on the comparison of two groups and the research interest is in the difference in means, rather

than variance. Cohen's tables on the T-test for means were used (p29). Using tables listing the values of the significance criterion, the effect size to be detected and the desired power in tables 4.4.1, the necessary sample size can be found. The effect size suggested for this study is 0.4, a moderate size typical of educational research (Hattie, Biggs, & Purdey 1996). To achieve 80% power with an effect size of 0.4 it is necessary to have a sample of 120. It should therefore be possible for the AGI study, with a sample of 132 to achieve this confidence level.

Accordingly, one hundred and thirty two children were randomised between a groupskills intervention (AGI) and a curriculum studies group (CSG), running in parallel, eight children in each group. This was repeated in three consecutive terms and to avoid confounding intervention status with a school effect, both types of group occurred within each school. At the time of selection there were no significant differences between the problem scores of the two intervention groups, from teacher, parent or self-reports. By one-year follow-up most children had changed school and all had changed teachers. As a consequence, at follow-up teachers were blind to intervention status. In addition, since each cohort was identified a term prior to intervention, a 'waiting-list' control was available. As a consequence of initial funding problems the waiting list data comprises teacher data alone, and omits the first cohort. Over the course of the intervention and follow-up, either because of leaving the area, or because parents declined further involvement, contact was lost with seventeen children.

4.2.2: Data Collection Instruments

The AGI Study used three different classes of data collection instruments: psychometric tests, classroom observation and school-based performance data.

Psychometric Instruments are standardised instruments for assessing people's ability aptitude, attitude or personality. They are standardised in that they have fixed items and rules about how they are administered. Also the performance of a standardisation sample of people is known. The results of any individual can be compared with this standardisation sample. The AGI study uses these instruments completed by teachers and parents about children in the sample, and two forms completed by the children themselves. The core of this assessment battery are the instruments developed and

standardised by Achenbach, namely the Teacher Report Form (TRF) (Achenbach 1991c), Child Behaviour Checklist (CBCL) (Achenbach 1991b) and Youth Self Report (YSR) (Achenbach 1991d) and these are supplemented by a self-concept measure: the Multi-Dimensional Self-Concept Scale developed by Bracken and designed to measure 6 aspects of self-concept and completed by the child (Bracken 1992). The principle involved in completing the YSR, CBCL and TRF forms is that observations of respondents are used to build a profile of the child's problem behaviours, graded for severity and then clustered into syndromes. The MSCS differs in that it clusters items scores on a sliding scale around positive dimensions of self-concept. The forms are administered over time and changes noted.

The Achenbach assessment battery consists of the CBC, YSR and TRF, and these three although developed and normed in the United States, have extensive use in the UK and are used in clinical child and adolescent mental health settings as well as being amongst the most frequently used research instruments in child psychology and psychiatry: recent studies have quoted the use of the TRF and CBCL in a number of research settings and both have been extensively used in ADHD research (Barkley et al. 2004). The Achenbach battery has been criticised for having a format that is sufficient for identifying children at elevated risk for psychopathology, but less suited to yielding information that is sufficient for determining clinical diagnostic status (Halperin & McKay 1998) and the nature of its questions and the cut-off points that are determined by applying T-score criteria do make it a challenge to use in sub-clinical settings. It is also noteworthy that problem behaviour checklists often exclude behaviours that are rare in the population or that have very low base rates of occurrence because their inclusion compromises psychometric sufficiency (i.e., reduces internal consistency) (Myers & Winters 2002). Although extensively normed, the test-retest correlations are not reported over comparable periods for each of the instruments but this is really worryingly low in the case of the YSR (0.56). The inter-rater reliability is between 0.6 and 0.75 for TRF and CBC and not relevant or possible for self report (Achenbach 1991d). These figures are really not that impressive for instruments so widely used and with such status in the field of child and adolescent mental health. There is also a question about the use of Achenbach to establish clinical cut-offs: these are set at one standard deviation, or a t-score of 60. This amounts to 17% of the population, but because the correlation between the instruments is so low (Achenbach 1991a), this 17% will contain different children

depending on which instrument is used to identify the clinical population, with the effect of enlarging the clinical population to something approaching 25%, a figure not supported by epidemiological studies (see Chapter 1).

Despite these criticisms, the CBC, TRF and YSR are accepted instruments for gathering problem behaviour data that can be easily compared through complementary items in the battery across contexts and informants

Child Behaviour Checklist /4-18 (CBCL)

Parents and parent surrogates are typically among the most important sources of data about children's competencies and problems. They usually are the most knowledgeable about their child's behaviour across time and situations. Furthermore, parental involvement is required in the evaluation of most children, and parents' views of their children's behaviour are often crucial in determining what will be done about the behaviour. Parents' reports should therefore be obtained in the assessment of children's competencies and problems whenever possible. The CBCL is a widely used measure of parent-rated child disturbance and present a standardised report on children's adaptive functioning, and emotional and behavioural problems reported from the evidence of the previous months. The questionnaire presents an introductory section of 'standardised tests' to assess children's competence with respect to cognitive ability and knowledge of academic subject matter. (Achenbach 1991b), p 8). This section was not used in the current study, since the prosocial aspects of behaviour are more adequately recorded in the MSCS. The Achenbach prosocial scale (present on the CBCL, TRF, YSR) was not used to ensure cognitive equivalence of the intervention groups, this being adequately covered by the randomisation procedure. The prosocial section is followed by a more detailed section of 118 problem items that are scored 0 = Not True (as far as you know), 1 = Somewhat or Sometimes True and 2 = Very True or Often True. The CBCL has mean test-retest correlations of 0.93 over 1 week and 0.86 over 2 years. In addition, the mean inter-rater correlation was 0.76 (Achenbach 1991b). Responses to these items are scored on syndromes of Withdrawn, Somatic complaints, Anxious/Depressed, Social Problems, Thought problems, Attention Problems, Delinquent Behaviour and Aggressive Behaviour. These syndromes are aggregated into broadband scales of Internalising Problems, Externalising Problems and Total Problems. The CBCL is designed for practical uses where decisions are needed in particular cases and research applications that aim to establish principles that are generalisable and testable: "...the CBCL is

intended to utilise the fruits of research to improve practical assessment and to enrich research by linking it to practical assessment procedures” (Achenbach 1991a p162). Use of the CBCL has become widespread in the research and clinical communities: a paper published by Achenbach in 1991 detailed hundreds of published studies reporting correlates of the profile scales (Achenbach 1995). The standardising process transforms raw scores into T-scores, approximately based on normalised scores: a T-score of 50 represents the mean score for a population, 60 (1 standard deviation) represents the 'borderline clinical' cut-point. The data produced are presented as useful in school contexts, indeed common sense would agree that parents' reports are important in supporting school-based reports in a number of areas such as justifying eligibility for special education and learning support services. For the purposes of this research programme, having parents' report in a similar format to the TRF for teachers, the CBCL gives cross-informant data that can be easily compared.

Youth Self Report

The Youth Self-Report (YSR) is derived from the CBCL /4-18 to assess self reported problem behaviours in adolescents of 11 to 18 years of age. This questionnaire is standardised in the same way as the CBCL and divided into two sections, a competence (17 items) and a problem section. The competence section consists of 'standardised tests' to assess children's competence with respect to cognitive ability and knowledge of academic subject matter that, for reasons similar to those discussed in relation to the CBC, was not used. The problem section consists of 102 specific items covering a broad range of problem behaviours, 1 open-ended item for additional physical problems without known medical cause and 16 socially desirable items. The latter are included to replace CBCL 4-18 items considered to be inappropriate for adolescents. Problem items are scored 0 = Not True (as far as you know), 1 = Somewhat or Sometimes True and 2 = Very True or Often True. The YSR has mean test-retest correlations of 0.79 over 1 week and 0.56 over 7 months. Responses to these items are scored on syndromes of Withdrawn, Somatic complaints, Anxious/Depressed, Social Problems, Thought problems, Attention Problems, Delinquent Behaviour and Aggressive Behaviour. These syndromes are aggregated into broadband scales of Internalising Problems, Externalising Problems and Total Problems. Raw scores are usually transformed into 't'-scores, approximately based on normalised scores: a T-score of 50 represents the mean

score for a population, 60 (1 standard deviation) represents the 'borderline clinical' cut-point.

Teacher Report Form

The Teacher Report Form (TRF) is a widely used questionnaire in child and youth mental health clinical practice and research (Achenbach 1991) generating data across areas of positive as well as negative adjustment. This questionnaire is standardised in the same way as the CBCL and divided into two sections, a competence (17 items) and a problem section. The competence section consists of 'standardised tests' to assess children's competence with respect to cognitive ability and knowledge of academic subject matter that, for reasons similar to those discussed in relation to the CBC, was not used. Problem items are scored 0 = Not True (as far as you know), 1 = Somewhat or Sometimes True and 2 = Very True or Often True. The TRF has mean test-retest correlations of 0.75 over 2 months and 0.66 over 4 months. In addition, the mean inter-rater correlation was 0.60 (Achenbach 1991c). Responses to 113 individual items are scored on syndromes of Withdrawn, Somatic complaints, Anxious/Depressed, Social Problems, Thought problems, Attention Problems, Delinquent Behaviour and Aggressive Behaviour. These syndromes are aggregated into broadband scales of Internalising Problems, Externalising Problems and Total Problems. Raw scores are usually transformed into T-scores, approximately based on normalised scores: a T-score of 50 represents the mean score for a population, 60 (1 standard deviation) represents the 'borderline clinical' cut-point. This questionnaire measured teacher identified school-based behavioural change.

Multi Dimensional Self Concept Scale

The Achenbach instruments were complemented by the Multidimensional Self Concept Scale (MSCS), a further self-rating scale added in order to evaluate changes in self-perceptions relating to important areas of school and family life. It consists of 150 items, grouped into 6 sub-scales of high reliability (Cronbach's alpha for the total scale 0.98, and 0.85 or above for each subscale) (Bracken 1992). Test-retest reliability is 0.90. The development of the MSCS was based on three related goals, including the desire to create a psychometrically superior, more comprehensive and more practically useful self-concept scale that had been available previously. The measure of the success of

Bracken in achieving these goals is the use of the MSCS in practice and research that is indicated through published studies. (Bracken et al. 1990). The Bracken questionnaire has less of a profile in the UK than the Achenbach battery and having been normed with a cross-section of American youth its relevance to the UK population might be questioned. However, the reliability data provided (Bracken 1992) demonstrate the (largely successful) efforts made to show equivalence with a broad range of scales but in particular the widely used Self Description Questionnaire (Marsh 1990).

The MSCS scale itself consists of 120 items that are organised in the following sub-scales: Social, Competence, Affect, Academic, Family, Physical which added together give a Total Scale score. Each Item is rated as Strongly Agree (SA), Agree (A), Disagree (D), or Strongly Disagree (SD). Each sub-scale contains an equal number of positively worded items and negatively worded items. Positively worded items are scored SA=4, A=3, D=2, SD=1. Items with negative connotations (i.e. items that if endorsed signify poorer self concept) are scored in reverse order: SA=1, A=2, D=3, SD=4. The MSCS presents a scale that is thoughtfully constructed with items developed in keeping with the six sub-scales emerging consistently in the literature (Bracken 1992, p13) and items written to sample relevant behaviours and situations comprehensively within each contextual domain. Items were then road tested before being standardised using a sample of 2,501 children attending school aged between 9 and 19 years, in a group representing a geographically and ethnically diverse backgrounds.

Observation, ethnography & natural behaviour measures were used to balance the psychometric testing, which can be seen as containing contrived measures unnatural to those completing the forms, and likely in themselves to encourage the reporting of atypical behaviour out of the context of the behavioural whole. Observation was therefore used to gather additional data in the AGI study. Observation is a process we are all consciously involved in, and all psychological research, including both quantitative and qualitative methods involves at least some element of observation; this may be something as simple as reading a dial on a machine or as complex as observing group interactions. It is often difficult to separate out observation as a distinct method as it is intertwined and permeates all psychological methods. In the AGI study, the observation demonstrates a commitment to represent the behaviour of children in terms that are relevant to the real-world setting of classrooms. Although falling short of becoming a

phenomenological measure concerned with the experience of those concerned, the difference of approach is justified as providing triangulation with the other instruments; particularly the self reports and therefore increases the objectivity of the reported results.

The advantages of observation as a method are that because it is not intrusive and the observer is not reactive to the situation, it provides a picture of what is going on, who is involved, when and where things happen. It can illuminate processes and it can examine causality, suggesting why things happen as they do in particular settings. It can give access to phenomena that are often obscured (e.g. non-verbal cues) or not amenable to experimentation: those situations such as the classroom that do not recur in the laboratory. The chronology of events can be taken into account, and continuities over time can be looked at. There can often be problems with validity with direct observation protocols, depending as they do on subjective views and subject to biases of opinion and interpretation that cannot always be made explicit. The aspects of a situation that are being observed have to be carefully pointed out, as does a consideration of whether the observations of behaviour actually relate to the constructs under consideration. The AGI study uses a protocol developed by the research team in collaboration with Gateshead Psychological Services for use in the classroom. This protocol consisted of an on-task observation, a disruption tally, and a running record of class activity and was applied in the classroom at each data gathering point. No observation was considered possible of the intervention sessions. An observer in the small group would not be as unobtrusive as in the class group, but other reasons are to do with the volume of data produced by qualitative measures of this kind and the constraints of resource in delivering the study. The chief concerns are about generalisability of behaviour gains, i.e. do gains made in the intervention context translate into behaviours noticeable to teachers, parents and through observation in the classroom situation?

Direct Observation Protocol (DOP)

The recognised need for an ethnographic measure of adjustment to complement the standardised questionnaire data meant led the research team to consider what form this data should take. The classroom presented itself as a natural forum for observation and rich as a potential source of information about how the intervention affected actual observed behaviours in this key arena. These data would effectively complement information on reported behaviours provided by the questionnaires. The research idea in

engaging this form of observation was to compare the AGI with the CSG groups in terms of their classroom behaviour, and also to compare the AGI and the CSG groups with the behaviours of the rest of the class. Turning these research ideas into a set of procedures for generating empirical data was a challenge, not just in the design of an observational system, but also in deciding the aspects of behaviour to be studied within the reference of the overall research design, and then defining these aspects into variables. Croll states:

‘Decisions about research design sampling and data analysis are interdependent and it is important that the researcher should have a clear idea about the form and purpose of his analysis, the questions the research is designed to answer and what will count as answers before making decisions about who and what to observe by what procedures and in what circumstances...Like all research designs, the adequacy of an observational system can only be judged with reference to its purposes. It is impossible to evaluate an observational system without reference to the purpose it is intended to achieve.’

(Croll 1986) p49).

Croll goes on to argue that it is because empirical data gathering procedures have their validity only with reference to a specific investigatory purpose, researchers should not be afraid to devise new observational systems. Underlying this argument is the recognition that it would be foolish to think that there is one true description of what is happening in classrooms and more obviously truthful to recognise that all procedures “involve abstracting from the infinity of things which could be said about any social setting those which are relevant for a particular investigatory purpose” (p50). Using the expertise of the Gateshead Schools Psychological Service, the measure for direct observation was developed to act as a measure of the impact of the interventions on classroom behaviour. The DOP was designed to provide data suitable for quantitative analysis and used a simple category system, making tally marks for behaviours sampled at regular time intervals. The model used for this type of observation was the Flanders Interaction Analysis Categories (FIAC). Flanders (1970) analysed verbal interaction between teachers and pupils, and although outdated in his positivist outlook, his widely used system allows complex classroom interactions to be quantified and therefore more easily analysed. The category system notes behaviours every time they occur, making an

attempt to show all interactions and preserving the flavour of the learning session. The FIAC system involves coding what is happening every three seconds. Other important models include the Exeter schedule as developed by Wragg (1994 p43). This schedule uses a sign system tallying each category once no matter how many times it occurs within each of the time segments. This schedule is in two parts: the first deals with the pupils' misbehaviour and how the teacher does or does not respond to it; the second consists of an individual pupil study which allows the observer to build up a profile of how involved in the task children appear to be and whether or not they misbehave. No standardised data are available for these measures, and it is possible that this represents difficulties in standardising observations that are by their very nature subjective and open to extraneous variables. This having been said, the DOP sampled behaviours often and long enough to rule out random variations due to the phase of the lesson, the time of day, the weather, teacher mood and lesson content. Protocols like this are widely used and variously reported (Bronson 1994; Reid & Maag 1994) giving some considerable confidence that despite the limitations, this form of observation remains a consistent tool for gathering data about the nature of children's behaviour.

On Task Observation.

The On Task Observation registers five categories of behaviour among the children in the class involved in the intervention study (i.e. those receiving either the Action Group Skills Intervention or the Curriculum Studies Condition). The Five categories are On –Task, Off-Task, Disruptive, Waiting and Other (marked as prosocial or withdrawn). The children are identified to the observer through a seating plan. The observer counts to four, looks up and locates child number one before marking his/her behaviour. It should take 2 minutes to score 16 children (the maximum number of intervention children in a primary class).

The Disruption Tally.

The Disruption Tally is a roving observation of the behaviours of all children in the classroom. Using a map of the classroom layout, the observer records behaviours in five categories as they occur and marks them beside the point on the map where they occur. Children are relatively static in observed lessons and therefore the points on the map relate to individual children. The five categories are Helping, Talking out of turn, Hindering other children, Out of seat and Aggression verbal/physical. Aggression was

found to be reported by observers so rarely so as to provide no really meaningful data. The ratio of observed non-aggressive problem behaviours Group: Rest of Class is derived by dividing the average value of observed non-aggressive problem behaviours for the intervention groups (AGI & CSG for hypothesis 5 & 6 and AGI or CSG for hypotheses 7 & 8) by the average value of observed non-aggressive problem behaviours for the rest of the class. When both values are equal the score is 1. When varying between 0 and 1 the observed behaviours of the group are less than those observed in the rest of the class. When varying between 1 and infinity the observed behaviours of the intervention groups are greater than those observed in the rest of the class. A value of 0.5 will indicate that average observed behaviours in the intervention groups were half those observed in the rest of the class, and a value of 2 will indicate that the observed behaviours in the intervention groups are double those observed in class.

The intervention children are observed for four hour-long sessions at each data gathering point. Two of these hours are in the morning on different days; two are in the afternoon on different days (an afternoon session may follow a morning session on the same day). In this way some account is made for mood and extraordinary conditions. The hour session is divided into four 15-minute periods and in each period the children's behaviour is observed on all three measures according to the following pattern:

On-task observation.	2 minutes
Disruption tally	6 minutes
On-task observation:	2 minutes
Running record	5 minutes

Observation One is therefore conducted eight times in an hour and 32 times at each data collection point. In marking behaviours emphasis is placed on the quantity of information being of value, the consistency of observer response and the instantaneous nature of the observation.

The scores for behaviours are then entered onto a collation sheet that gives totals for the hour, for the morning and the afternoon, and for the entire data collection point. This

gives quantitative data for observations one and two that can be entered into an analysis package to be compared at each time point.

The aim of structuring the data in this way was to define the variables of behaviour that inform the wider research design and to locate them in time and locate them in the context of what is happening in the classroom. The aim was to provide data for empirical and comparative analysis. Establishing reliability is therefore a key concern. The main difficulty, however, in establishing data reliability is the conceptual problem of deciding what is meant by reliability in this context. The idea of reliability in the statistical sense is strongly associated with establishing the reliability of measurement procedures in attitude scaling and for tests of achievement or ability. In this sense the DOP suffers from the direct comparisons that will be made with the CBCL, TRF, YSR and MSCS. Reliability in this sense is concerned with the replicability of the scores established by a test procedure and the interrelationship of the items in a measuring instrument, and is established by demonstrating that the measurement instrument will establish the same result on a second occasion when that which is being observed is unchanged (test-retest reliability) or else that different combinations of items in the test battery, taken separately will arrive at the same result. According to Croll the application of this model “causes some confusion in the analysis of observational data where the question of the reliability of the observation instrument as used by different observers is of paramount importance and where there is some dispute over whether the ability to discriminate consistently between different objects of observation should be a criterion for a satisfactory observation instrument” (Croll 1986 p150). Croll goes on to describe how some researchers use the terminology ‘Observer Agreement Coefficient’ to refer to the extent to which different observers code the same observed phenomena in an identical fashion and ‘Reliability Coefficient’ to refer to the extent to which individuals are dependably discriminated. According to Croll “the distinction between the two types of analysis is a fundamental one” (p150), and provides the basis for his justification of the reliability of systematic observation measures. Using these two aspects of reliability to reflect on the DOP as used in the AGI research is instructive. In the first case, the emphasis on observer agreement as of central importance is recognised, as is the practice of showing that the procedures can be used by different observers to arrive at identical descriptions of the same events. Although the DOP was piloted by two members of the research team, data were never gathered to establish this Observer Agreement Coefficient.

Similarly, the second aspect of reliability is recognised as important, this aspect refers not to the agreement between observers on the use of categories of a variable, but to whether the variables can consistently distinguish between individuals or classrooms that are observed. However, no data relating to the performance of the DOP in the AGI research design were gathered in relation to this reliability coefficient. Although Croll does state that "This is a more problematic application of the concept of reliability to classroom observation data and it is by no means clear that it is necessary for all observational data to be reliable in this sense" (p152), it is an omission in the AGI design not to have data to help in the clarification of these issues. The reasons were in part due to the time constraints in gathering the first batch of data, and in part due to the confidence placed in the experience of the Gateshead Psychological Service in the construction of the DOP from established elements of effective observational practice see e.g. (Keeves 1988), (Edwards & Westgate 1987). In the final analysis the reliability of the data must rest on this confidence in a service with great experience in gathering data of this type in the schools of the North East region, although the question mark over the reliability of the observation data will affect the interpretation of the data as it is analysed.

Individual and Group Attendance Records

School-based attendance data on children in the AGI study was taken from the logs kept by schools. School attendance records are kept as part of the legal school records. A register is marked at the beginning of the morning and afternoon sessions with children gathering in from classes before moving off to the lessons. The school tracks children who are persistent poor attenders and monitors truancy in this way. The study uses these school records to measure the impact of behaviour gains on school attendance. These are epidemiological data and as such stable. However there is a problem in using the data to demonstrate the impact of a groupwork programme that lasts a mere 12 weeks. There are many factors in school non-attendance: for example there is a marked decline were children change school, an extraneous variable that is difficult to compensate for. Because a large proportion of the research population changed school over the period of the follow-up it is difficult to compare from a baseline. Typically there is an expected increase in non-attendance as children move to secondary school as the controls that arise from a smaller student population in primary and middle schools are

no longer active in the same way. The result is that the data begin to reflect the school environment rather than the individual behaviours of the children.

4.2.3: Data Collection Procedures

Sampling Issues:

The way in which cases are selected for research has a part to play in the overall argument that leads to research conclusions. If the research is concerned with providing information about a single case – the treatment of a single patient, the evaluation of a single experimental institution and whether it meets its aims, for example – then case selection is obviously unproblematic. More often the research will be making claims about a population, a class or group or kind of person etc. When this is so, a part of the argument in a research paper entails showing that the cases investigated can be taken as typical or representative of this population.

Making a sample survey involves making implicit claims to representation – that the person, school or city location is typical of others, and therefore that what is said about it can be applied in other contexts. There are limits to the guarantee of representativeness: it is difficult to be assured that the sample represents the attributes in the same proportion as are found in the population. However every attempt is made to establish that the sample is typical, that other cases resemble it enough for general conclusions to be drawn. A part of the assessment of a research design is the deployment on behalf of the reader of an analytic imagination in trying to see ways in which the cases selected might not be typical and the conclusions drawn from them subject to qualification.

In the case of the present study into the effects of groupwork, children are selected by teachers on the basis of the presence of anti-social behaviour patterns. The groups are constituted with equal numbers of boys and girls including a range of severity of presenting symptoms. These groups are intended to contain a representative sample of the types and severity of anti-social behaviours present in ordinary schools where disaffected behaviours affect up to half of a class population. The targeted behaviours were those that existed in a form severe enough to cause concern to teachers or parents, and to impede scholastic progress and social integration. These children were selected from primary, middle and secondary schools in the North East of England, and

the groupwork developed in that specific context. It is argued that the children included in the research sample are typical of the kinds of schoolchildren who display anti-social behaviour of concern to their teachers and parents, and the logic of impacting upon policy will depend upon any benefits from groupwork generalising to other children in other geographical and cultural circumstance.

Table 4.2.3: Timetable of Intervention and Data Collection

	1996			1997			1998		
	Jan-Mar	Apr-Jul	Aug-Dec	Jan-Mar	Apr-Jul	Aug-Dec	Jan-Mar	Apr-Jul	Aug-Dec
Time 1 Waiting List	2A 2B 2C	3A 3C	4A 4B						
Time 2 Immediately Prior Interv	1A 1B	2A 2B 2C	3A 3C	4A 4B					
Intervention	1A 1B	2A 2B 2C	3A 3C	4A 4B					
Time 3 Immediately Post Interv		1A 1B	2A 2B 2C	3A 3C	4A 4B				
Time 4 12m Post					1A 1B	2A 2B 2C	3A 3C	4A 4B	

Data Collection Procedure

Data collection was co-ordinated by a Research Associate, paid part-time, who delivered a strategy of collection decided in discussion with the Research Leader who also delivered the interventions. At each collection point a group of Research Assistants were employed for a month to help with the gathering of data. The Research Associate was a graduate who had been working on other university-funded programmes and was in the process of completing a degree in psychology through distance learning. The Research Assistants were from a range of backgrounds, either undergraduates, researchers from other projects or retired teachers.

The Teacher Report Forms (TRF) were completed by teachers who had some knowledge of the child outside of curriculum delivery classes i.e. as classteachers or tutors. In the first wave of assessments this was the class teacher in the primary schools, the form tutor in the middle schools and secondary school. Ensuring completion

with these staff was straightforward, as they were the school link with the research programme and the majority were sympathetic to the aims of the research. In this way the forms, although tiresome for teachers to complete, were treated with the respect that the project deserved. In subsequent waves, the criteria for choosing a teacher to complete the form remained the same, but these teachers were different, as the children in the project changed year groups and in the case of the primary children, changed schools. The staff at follow-up points knew substantially less about the programme of intervention and research, were a more disparate and numerous group, and were less enthusiastic about completing the forms. In order to overcome these factors, a financial incentive was offered to teachers in the form of payment per script; something teachers are familiar with from the marking of exam papers. In addition, the teachers concerned all received a letter explaining the aims and function of the research and intervention programmes and in most cases a visit from the research associate to underline the personal investment of the team in getting information back about children.

The Youth Self Reports (YSR) and Multi-Dimensional Self –Concept Scale (MSCS) were completed by children in school time through the means of a specially convened session, formally conducted by retired teachers working for sessional payment. The entire intervention group was gathered at the same time (AGI+CSG) and both measures were completed in the session. The children were not especially enthusiastic about these sessions and a small minority (about 5%) expressed reluctance and questioned the importance of form filling. The sessions were therefore started with a brief summary of the programme aims and the importance of quality data in audience-appropriate language. Also chocolate rewards were used as an incentive for completion.

The Child Behaviour Checklist (CBCL) were the most problematic forms to pursue and it is no surprise that the organisational complexity in managing a dataset including parent measures is daunting for research teams and explains the relative lack of parent data in studies of this type. Parents were informed of the research aims and assessment protocol in appropriate language during the initial consent procedure. At the data collection points CBCL forms were mailed to the parents of the children involved in intervention using the school data on home addresses. The CBCL form was accompanied with a letter warmly thanking the parent for their efforts in completing the form and a stamped self-addressed envelope for the return. About 60% of the forms

were returned in this way at time T1 falling off to 50% at T4. For those parents who did not return the form in this way a telephone call was made to the home, again using school data, and a pleasant reminder was made, along with the offer of a home visit to collect the form. Often messages were left with relatives and answer phones. With still no response, a further telephone call was made with a gentle reminder and a request that a visit be made to pick up the form. These visits were made by research assistants (RA) working in pairs for the main part and alone when unavoidable. During the visit the RA would sit with the parent or carer and ask the questions on the form to the parent while marking the answers on a new form. If there was nobody in when the RA called, a card was left telling the parent when the next visit would be made. With the card was a message that if the parent no longer wished to take part in the study, then please would they make contact with the research team or the school. After three visits with no contact the efforts were concluded. There was a definite effect of time, parents becoming more reluctant to fill in forms at time 4. Overall the return rate was 65-70% as the attrition figures presented later will confirm.

The Direct Observation Protocol was conducted by Research Assistants in school in school time. In primary schools completion was the most straightforward as all the intervention children were from the same class and could be observed in the four sessions described above. In the middle and secondary schools, the intervention children did not ever come together in the same classroom and the protocol had therefore to run in a structure of great complexity in different lessons to ensure that each child was observed for four hours. The RA would work closely with the school and in particular the head of year in determining which classes would provide the optimum number of individual assessments before gaining the permission of the class teacher for their attendance in the lesson. At all times it was stressed that the observation was on the children, and no data were collected on teaching styles. A pool of students and research assistants were used to collect this data, all of whom were blind to treatment status, but there exists the possibility of an observer effect depending on attitude to the project – it must be stressed however, that the observers were all undergraduates or postgraduates, were given adequate training, and tutored in the skill of impartial observation. All appeared to be keen and sympathetic. The children were kept blind to the function of the observer in the classroom, and the connection of the observer to the

group intervention and subsequent research by the nature of the introduction given by the classteacher and by the use of different observers at different time points.

Attendance Data were collected by RAs working on school sites from school records with the permission of the Head teacher and the parents.

Section 4.3: Data Analysis Procedures

4.3.1: Reporting Conventions

T-Scores or Raw Scores:

Reported below are the T-score calculations made on raw scores from the YSR, MSCS, TRF and CBC. T-scores are favoured because they are standardised against a normal population and take account of the age and sex of the subjects they report. Declining T-scores means improvement and a score of 50 represents the mean score for a population, and 60 (one standard deviation) represents the 'borderline clinical' cut-point.

Reporting Analyses

The AGI study adheres wherever possible to the principles developed by John Hattie to improve the quality of statistical studies (Hattie & Flounders 2004). These principles are echoed elsewhere in statistical publication (Hopkins 2002), and adopted because of their up to date engagement of the literature, their pragmatic approach, their scholarship and their particular relevance to the educational setting. Several of the principles have a direct effect on the reporting of statistical analysis in the AGI study relating to the articulation of the limitations and place of statistical significance testing that is summed up in this quote:

"If we can control statistical significance simply by changing sample size, if statistical significance is not equivalent to scientific method, and if it has only questionable relevance to one out of fifteen threats to research validity then I believe we should eliminate statistical significance testing in our research." (Carver 1978 p392).

This might be regarded as something of an extreme position, and the author was to revisit the subject with the perspective of 15 years experience (Carver 1993), but out of

the broad thrust of this approach, the meaning of statistical significance was kept firmly in mind during the reporting of AGI results: statistical significance means statistical rareness. Results are statistically significant because they occur very rarely in random sampling under the condition of the null hypothesis. Hattie does not suggest that a hypothesis test should not be performed, but rather that when authors employ statistical hypothesis testing they should:

- a) Use the term 'statistically significant;' instead of 'significant' when interpreting their results.
- b) Not to become bound by arbitrary rules (such as $p < 0.05$, $p < 0.01$), but rather to paint a defensible interpretation of their results by using patterns across many substantively important findings.
- c) Accompany p-values with an index of the size of the effect or strength of association
- d) Search for replicated evidence. Finding similar effects across many studies is more convincing than finding statistical significance in any one study.

(Hattie & Fiandaca 2004)

Wherever possible, these principles have been followed in the reporting of results in the AGI study.

4.3.2: Analysis and Effect Size

Analysis:

T-scores were analysed using a repeated measures analysis of variance (ANOVA) to identify change over time. Total scores are reported along with the sub-scales of Internalising and Externalising (YSR, TRF, CBC). The sub-scales of Academic, Affect, Competence, Family, Physical, and Social are reported for the MSCS along with the total.

The data were first tested for sphericity using Mauchly's test of Sphericity, a test that examines the data for the criteria of regularity and variability necessary for running an ANOVA test. Mauchly's Test of Sphericity gives a significance level, and if this level is above 0.05 then sphericity is assumed. If the significance level is below 0.05 then the data are corrected so that the criteria of variability and structure can be maintained.

These Epsilon corrections can take three forms depending on the degree of conservatism required: all make a manipulation of the Degrees of Freedom (df) used in calculation of the ANOVA by suggesting a figure that when multiplied by the existing df gives a new corrected value for use in the ANOVA. The Epsilon correction used in this analysis was the Greenhouse-Geisser, because the two alternatives: the Huynh-Feldt and Lower Bound, were considered too inflexible as their value does not change, but merely define the higher and lower boundaries of conservatism for the calculation.

The full report of findings found in the appendix gives figures for the sum of squares, the degrees of freedom, the mean of squares and the F-value calculated from the sum of squares which when combined with the degrees of freedom allows the significance value to be reported.

Effect Sizes:

Effect size is a way of quantifying the effectiveness of a particular intervention, relative to some comparison. By placing the emphasis on the most important aspect of an intervention – the size of an effect – rather than its statistical significance (which conflates effect size and sample size), it is argued (Coe 2000) that Effect size promotes a more scientific approach to the accumulation of knowledge. The ANOVA test is a rigorous test to determine whether the variance between groups is greater than variance within groups ('groups' here can be either the same population at different time points, or different populations), and it delivers a p-value that is a reflection of the size of the sample as much as it is a measure of the size of the effect. Effect size, on the other hand is a measure of magnitude of change and the analysis quantifies that change in a way that is independent of sample size. The effect size calculation takes into account the distribution of the populations under comparison by using the standard deviation at each comparison point. This measure of degree of change uses a calculation based on the formula $[m_1 - m_2 + \sqrt{(sd_1^2 + sd_2^2)}]$ where m_1 is the mean score at time 1 and m_2 is the mean score at time 2, sd is the standard deviation at times 1 and 2. (Becker 2000). The effect size gives a value for the number of standard deviations by which the mean score changes between assessment points. One standard deviation = one unit of change = an effect size of one.

In hypotheses 1 and 2 the formula $[m1-m2 + \sqrt{(sd1 \times sd2)}]$ is used in that form where $m1$ is the initial mean score of the entire subject population and $m2$ is the subsequent mean score. The subtraction is done this way round because declining scores mean improvement that is predicted as a result of intervention, so the subsequent score is predicted to be smaller than the initial score. A positive effect size will therefore indicate change in the direction predicted by the hypothesis.

Hypothesis 1 is concerned with the change that occurs to the whole subject population over time. There are three time points and effect sizes are calculated for the time periods T1-T2, T2-T3, T1-T3.

Hypotheses 2 uses the same population over three time points: T2, T3 and T4. Effect sizes are calculated for the periods T2-T3, T3-T4 and T2-T4.

In Hypothesis 3 and 4, the Action Group Skills Intervention (AGI) is compared to the Curriculum Studies Group (CSG) and the formula is modified to:

$(AGIm1-AGIm2)-(CSGm1-CSGm2) + \sqrt{(\sqrt{(AGIsd1 \times AGIsd2 \times CSGsd1 \times CSGsd2)})}$.

Because the hypotheses predict positive change over time, the subsequent-scores are still subtracted from the initial scores. Further, because the hypotheses predicts positive change in AGI to be greater than CSG, the product of the curriculum studies group scores is subtracted from the product of the drama group scores to give a positive effect size score indicating change in the direction predicted by the hypotheses.

The reporting the effect sizes by ascribing values to categories such as "trivial" or "large" is the subject of some confusion (Hopkins W.G. 1997) and the modifications suggested by Hopkins are adopted in the AGI study. Cohen has suggested that the effect size statistic of 0.8 is large, 0.5 is moderate and 0.2 is small (Cohen 1988). Hopkins has argued that the process of making the independent variable normally distributed, then "dichotomizing" it by splitting its values down the middle to make two fitness groups (Cohen's method) throws away information and therefore Cohen's values for effect size underestimate the true magnitude of the effect (Hopkins 2000). Hopkins goes on to suggest an approach extrapolating the correlation coefficient scale, keeping the scale linear for correlations and frequency differences, and adopting a Likert-scale approach in assigning category names. The resulting table is the reference for those effect-size values reported below:

Table 4.3.2: Comparing Cohen's and Hopkins' effect size categories

Effect Size	Improvement in percentile rank for average member	Cohen's Categories	Hopkins
0.1	4	<0.2: trivial	<0.2: trivial
0.2	8	>0.2 and <0.5: small	>0.2 and <0.6: Small
0.3	12		
0.4	16		
0.5	19		
0.6	23	>0.5 and <0.8: moderate	>0.6 and <1.2: moderate
0.7	26		
0.8	29		
0.9	32	>0.8: large	>1.2 and <2.0: large
1.0	34		
2.0	48		>2.0 and <4.0: very large
3.0	50		
>4.0			Nearly perfect

In calculating the effect size the standard deviation is usually calculated from the sample of values available, and should be regarded as an estimate of the 'true' population value, especially where the sample is small. As a result the standard deviation is subject to sampling error. Although using the pooled standard deviation to calculate the effect size generally gives a better estimate than the control group SD (Hedges & Olkin 1985), it can be shown that the sample values are on average a little larger than the population from which the sample is taken. Ideally therefore a correction is used, especially if the sample is small. Hedges and Olkin give a formula which provides approximate correction (Hedges & Olkin 1985), but this was omitted from the AGI study calculations. With the sample size not excessively small this is not judged to be a serious omission.

Both Coe and Hattie (Coe 2000; Hattie, Biggs, & Purdey 1996) give examples of effect sizes from published research in different disciplines and there is some agreement that the effect size for educational research programmes in the area of delinquency and targeted intervention averages at 0.4 (Slavin & Madden 1989;, Lipsey 1992). Therefore the AGI will highlight as significant effect sizes of the order of 0.4.

4.3.3: Maximising Reliability

One or Two Tailed Testing:

The p-value is calculated for both tails of the distribution of the statistic. This follows naturally from the meaning of statistical significance and why tests are called two-tailed. In principle it is possible to eliminate one tail, double the area of the other tail, then declare statistical significance if the observed value fell within the one-tailed area. The result would be a one-tailed test. The Type 1 error would still be 5% but there is more power to detect the effect.

Usually it is argued that a one-tailed test is permissible only if there is good reason beforehand that the outcome will be positive. The counter-argument is that in a confidence interval view of significance, what constitutes 'good reason'? Is it not the case that there has to be absolute certainty of a positive value, which undermines the function of statistical testing? The correct view is unquestionably to place confidence limits equally on each side of the observed value (two-tailed testing).

To do otherwise is to interpret the p-value differently. P-values for one-tailed tests are half those for two-tailed tests. It follows that the p-value from a one-tailed test is the exact probability that the true value of the effect has the opposite sign to what is observed, and that $1 - p$ is the probability that the true value of the effect has the same sign. Under this interpretation it is possible to take the position that: "All tests of significance are one-tailed in the direction of the observed effect. The resulting p-values represent the probability that the true value of the effect is of the sign opposite to the observed value." (Hopkins, 2000 webpage). This position dodges the issue of predicting with certainty in a particular direction, making the one-tailed prediction post-hoc in essence. The advantage of this is that there is a greater chance of finding significance, the disadvantage is that the test is a weaker one.

It has been argued convincingly (Dracup 1995), that significance is altogether a more complex matter, and that the above is a misinterpretation of the basic statement that "the probability of a type 1 error equals the significance level". According to this view, if result is significant at a particular level, then the probability that the null hypothesis is true is equal to that significance level. Here significance level is being correctly interpreted as a conditional probability, but misinterpreted as the probability that the null hypothesis is

true given that the test has led to its rejection, rather than the probability that the test will lead to its rejection given that the null hypothesis is true. The wording is similar Dracup argues, but the meanings are very different: the erroneous interpretation gives a false sense of security about the validity of conclusions and leads to the disregard of two important factors in judging the results of research. These factors are a) the prior probability that the null hypothesis is true $p(H_0)$, and b) the power of the experiment $(1-\beta)$.

Dracup follows Pollard and Richardson (1987) in using a Bayesian analysis of the hypothesis testing process to illustrate the importance of these two factors. Bayes' theorem allows probabilities associated with hypotheses to be updated in the light of new evidence, stating that the probability that the null hypothesis is true given a significant result =

$$\frac{\alpha}{\alpha p(H_0) + (1-\beta)(1-p(H_0))} \times p(H_0)$$

The equation shows that the probability that the null hypothesis is true given that the result of the experiment is significant is a function of three factors: the significance level, α ; the prior possibility that the null hypothesis is true, $p(H_0)$; and the power of the test to detect a real experimental effect, $(1-\beta)$. This position is very different, it is argued, to the erroneous interpretation that the probability that the null hypothesis is true given that the result of the experiment is significant is equated with the significance level. The values of $(1-\beta)$ and $p(H_0)$ will not be known for any real experiment, so, as Pollard and Richardson argue, the probability that the null hypothesis is true given a significant experimental result is indeterminate.

This is worth mentioning because the fact that numerical values cannot be assigned to these two factors cannot be taken as a justification for ignoring them. Both factors are under the influence of the researcher and must be considered when judging the impact of a significant experimental result. Consideration of Chapter One will allow a judgement to be made as to whether due account has been taken of the literature in the selection of the hypothesised effect, because thorough reading of the literature lowers the prior probability $p(H_0)$ that the null hypothesis is true before the experiment is conducted. A

good experimenter, however, will not allow the probability to be lowered to zero. Similarly consideration of the method section would allow insight into the adequacy of the experimentation and whether the sample size takes adequate account of the likelihood of finding an effect (experimental power), and as Cohen (1992) has pointed out that tighter experimental control, through matching participants for example, has similar effects. Even when this examination is complete, it will not be possible to assign a numerical value to the probability that the null hypothesis is true given that a significant experimental result has led to its rejection. Nonetheless it is more likely that credit and credence will be given to the significant results of carefully planned and conducted experiments investigating hypothesis arising from a detailed consideration of existing literature.

In this way, although all the hypotheses below predict change in one direction, having good reason to do so from readings in the literature of intervention, all the tests on the data are two-tailed. The two-tailed test is stronger and with the larger sample that this study presents and the experimental power provided by the methodology gives a good basis for detecting experimental effect. The Bayesian analysis shows that the larger the power of the study (other things being equal) the smaller will be the probability that the null hypothesis is true given that the result of the experiment are significant. Provided the power of the study is greater than the significance level adopted (a two-tailed test is a good way of ensuring this) then the probability that the null hypothesis is true given a significant result will always be smaller than it was before the experiment was conducted

Multiple Testing:

There is a problem in evaluating a series of statistical tests as recognised by psychologists (e.g. Sakoda 1954). In essence the problem is that as the number of statistical tests performed increases, so does the probability of finding tests that display statistical significance. The general approach to the problem is to set the significance level (p) at .05 or .01 and to find the chance probability of obtaining at least n significant results. Sakoda (1954) has taken published tables of p at the .05 and .01 levels showing the probability of obtaining n or more significant statistics out of N calculated statistics and employing the normal curve approximation to the binomial distribution by calculating a critical ratio. Use of this normal curve approximation is only viable when $N=100$ or more when $p=.05$ or $N=500$ when $p=.01$. Plotting N on a logarithmic scale as completed

In the Sakoda paper means that it is possible to ascertain the likelihood of obtaining significance through multiple testing when the normal curve approximation is not applicable.

The Sakoda logarithms are used where multiple testing is an issue in the data presented below. This is of particular concern in Hypothesis 1b, 1(sub)b, 2b, 3b and 4b where the instrument of assessment yields 8 sub-scores for analysis across 3 time points.

Hypotheses 3 and 4 provide more statistics by looking at two populations. The issue of multiple testing is given attention in the results section by referring to the graphs presented in the Sakoda paper starting with the logarithms plotted for the .05 level of significance and repeated for the .01 logarithms. The number of calculated statistics N is 8×3 along the x axis. The number of statistics significant might be 5 (e.g.) and the intersection of these vertices lies above the curve labelled 0.01, which means the chance of finding 4 significant statistics by chance from 24 statistical tests is less than one in a hundred.

4.3.4; Hypothesis 9

Hypothesis 9 is something of an anomaly to the research conventions adopted in the AGI study and reported above. The emphasis given to the within-school conditions in Section 3.1.1. creates a need to examine the effects of the differing conditions on the research outcomes measured through processes given above and reported below. There is a need to exclude the possibility that the effort made by the school into accommodating the intervention and research processes was responsible for intervention gains. If this was the case, the school could rightly claim that the crucial ingredient of behaviour change was down to their effort, and not due to the effects of intervention. As well as the interest in excluding this possibility, there is data to be gathered on the value of putting differential effort into supporting schools in the organisational aspects of putting an intervention in place.

These within-school conditions were felt to be important to the smooth running of the groups, and worth emphasising to the school hierarchy during the set up process, but of course variety was encountered in how these group and school factors were adopted from school to school. In order to test how this variety affected behaviour outcome, Hypothesis 9 looks for change in the positive direction:

Positive behaviour outcome for both AGI and CSG groups will correlate significantly with the attitude of the school to the interventions and research process, measured by scoring within-school conditions.

The Hypothesis is split into three parts (a – c) each relating to a different dataset.

When schools are ranked for outcome using a gain score (T2-T4) from child self-report (YSR), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions

When schools are ranked for outcome using a gain score (T2-T4) from teacher report (TRF), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions

When schools are ranked for outcome using a gain score (T2-T4) from parent report (CBC), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions

In order to test this hypothesis, the children from each school were isolated in the dataset in turn. Gain scores relating to each school were calculated by subtracting individual scores at T4 from scores at T2. Since decreasing scores indicate behavioural gain, a positive gain score indicates an improvement over the period. First the scores for all the children from each school were calculated, and then those from each school receiving the AGI intervention were highlighted, followed by those from each school receiving the CSG intervention. Gain scores were calculated in this way using total T-scores for TRF, CBC and YSR forms.

Table 4.3.4a: TRF Gain Scores T2-T4

	Whole Group		AGI		CSG	
	mean	sd	mean	Sd	mean	Sd
STC	7.88	12.94	12.75	16.07	3.00	6.78
CEN	-3.04	8.03	-0.43	6.75	-5.85	8.59
MWP	11.55	10.64	11.14	8.13	11.93	12.82
JSP	14.00	12.48	16.54	10.68	12.00	13.78
PSJ	3.00	8.30	5.50	11.93	0.50	1.29

Table 4.3.4b: YSR Gain Scores T2-T4

	Whole Group		AGI		CSG	
	mean	sd	mean	Sd	mean	Sd
STC	1.62	16.24	5.33	13.62	-1.57	18.63
CEN	6.95	12.36	2.73	9.97	11.18	13.50
MWP	7.52	15.36	9.54	18.36	5.64	12.34
JSP	2.71	12.53	0.17	13.59	4.63	12.25
PSJ	7.43	13.84	3.25	9.74	13.00	18.74

Table 4.3.4c: CBC Gain Scores T2-T4

	Whole Group		AGI		CSG	
	mean	sd	mean	sd	mean	sd
STC	-7.44	6.80	-7.40	7.09	-7.50	7.51
CEN	-0.20	6.30	0.45	7.04	-1.00	5.57
MWP	4.95	10.52	4.71	13.62	5.07	9.20
JSP	-0.50	6.75	-7.00	8.48	2.75	3.10
PSJ	-3.14	5.93	-2.33	5.86	-3.75	6.80

Using these gain scores, schools were ranked according to which school hosted the greatest gain score for the TRF, YSR and CBC, with totals produced for the whole group, AGI and CSG conditions.

Following this, each school was given a score by the practitioner delivering the interventions from 1 to 10 on the following criteria: ease of getting appointment with Headteacher, suitability and commitment to intervention programme of co-ordinating teacher, responsibility taken for parental contact and responding to parental concerns, adherence to selection procedure, room suitability and privacy, ease of release of children from timetable, and commitment of classteacher to post-session feedback. These scores were added to give a total for within-school conditions and the schools ranked according tot hat total.

Table 4.3.4d: School Scored According to Commitment Criteria (1-10)

Criterion	JSP	CEN	MWP	PSJ	STC
Contact Headteacher	7	9	9	6	10
Co-ordinating Teacher	8	9	9	4	10
Parental contact	7	9	9	7	9
Selection	5	8	9	5	10
Room Suitability	8	6	7	7	7
Release from timetable	8	8	9	9	10
Feedback	4	8	8	6	9
Total	47	57	60	44	65
Rank Order	4	3	2	5	1

Although this scoring procedure was completed some time after the interventions had taken place, issues of bias were addressed through a scoring procedure that was completed before the calculation of the gain scores. In addition the criteria were scored across schools rather than by school, which goes some way to preventing manipulation of the rank order.

In order to analyse the these data and inform the hypothesis, the rank order in school commitment was correlated with the rank order of gain score using Spearman's rank order correlation. This analysis is a nonparametric (distribution-free) rank statistic proposed by Spearman in 1904 as a measure of the strength of the associations between two variables (Lehmann and D'Abrera 1998 pp292, 300), it can be applied to compare two independent random variables, each at several levels (which may be discrete or continuous). Spearman's method works by assigning a rank to each observation in each group separately, working on ranked (relative) data, rather than directly on the data itself, unlike the product moment correlation coefficient. A value of the coefficient near one indicates good agreement; a value near zero, poor agreement, generally, levels of 0.8 and above are considered as significant. Of course, as a distribution-free method, the Spearman rank correlation does not make any assumptions about the distribution of the underlying data and is a measure of monotone association that is used when the distribution of the data make Pearson's correlation coefficient. It is calculated using the formula:

$$r_s = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$$

where d is the difference between the ranks and n is the number of observations.

Section 4.4: Applied Ethics

Ethical issues are of central importance to any research design, and every new research study aims to redress the classic lines of investigation that raised serious questions about respect for participants to which the investigators did not pay sufficient attention. Professional associations have made steps in establishing a common ground for ethical reference when designing and conducting research, and these codes of conduct are readily accessible to researchers through the Internet: (British Psychological Society

2004; British Sociological Association 2004; Social Research Association 2004). Despite the wide circulation of these principles and their availability, the debate on ethical conduct of research is muted. Bryman suggests that debate is frustrated because of widely ranging opinions of what is acceptable, and the perception that the debate has not moved forward a great deal since the 1960s (Bryman 2001). On top of this, when attention is focused on the extreme and notorious cases of ethical violation, as seems to be the case, the debate is skewed towards the extreme, with the resulting perception that ethical issues reside predominantly in some methods (e.g. disguised observation) and others (ethnography for example) are immune from ethical problems. This will have an inevitable effect on the depth to which ethical principles are engaged.

Discussion about ethical principles in social research tend to revolve around issues that recur in different guises but that have been usefully broken down into four main areas:

- whether there is harm to participants
- whether there is lack of informed consent
- whether there is an invasion of privacy
- whether deception has been used.

See e.g. Diener & Crandall (1978).

While it is recognised that these areas are not fully discrete (it is difficult to imagine how, for example, the principle of informed consent could be built into an investigation in which research participants were deceived) they do provide a useful classification, and a starting point from which a stance on ethics can be made. Practitioners typically adopt a relationship to ethics, from which various stances can be distinguished: *universalism* – taking the view that ethical precepts should never be broken, *situation ethics* – which considers precepts on a case by case basis, *transgression is pervasive* – submitting to the notion that virtually all research contains elements that are at least questionable and finally *anything goes* – whereby flexibility is accepted in ethical decision making.

The second of these stances is of most interest to the AGI study, and the one that reflects most accurately the position adopted in the delivery of the methodology and the

completion of the research process. Situation ethics is more accurately described as principled relativism, the case for which is persuasively made by (Goode 1996). The position can be represented as accepting that the end justifies the means, often accepting also that there is little choice to engage in dissimulation on occasions if we want to investigate the issues in which we are interested (see e.g. Holdaway (1982). This position can be discerned in the specific research issues presented in relation to the AGI study. At the same time the British Psychological Society's statement on 'Ethical Principles for Conducting Research with Human Participants' (British Psychological Society 2004) is adhered to in a way that is as close as possible in order to complete the research. This statement begins by stating that investigators must consider the ethical implications and psychological consequences for the participants in a piece of research. The investigation must be seen from their standpoint, and foreseeable threats to psychological wellbeing, health, values or dignity should be eliminated. Protecting participants also encompasses invasion of privacy: if the research involves behaviour that a participant may regard as personal or private, they must be protected from stress. Observational studies must respect the privacy and wellbeing of those studied and if consent has not been obtained they should only be observed in situations where they might expect to be observed by strangers. Examples of current thinking in the area that have had an effect on the characterisation of the AGI position include Morgan, Harmon, & Gliner (2001) and Thompson (1990).

Ethical Issues In Subject Identification:

Using subjective criteria to identify 'at-risk' children is ethically problematic for the reasons that the perceived condition of risk exists, by definition, in a sub-acute form and may be overlooked by the significant figures in the child's life (including the child himself). Alternatively if the risk is acknowledged it may be attributed to passing circumstances of peer-group, development or family issues, and reasonable expectation expressed that the behaviour will pass according to time. The literature is, however, alive with evidence that problems identified at early stages in a child's developmental life persist into later childhood, often resulting in the involvement of specialist services. The National Child Development Study also has highlighted the trajectory of normative development and it has shown that for many children, behaviours present in early stages of development come to be seen as personality traits as the child matures (National Child Development Study. 2000). Those who function adequately in school and at home

find a way to incorporate these negative traits into socialised behaviour, but these children are often in a minority.

But an at-risk group must by definition stand a higher chance of containing a larger proportion of children who do go on to use services and therefore the at-risk group warrants attention. There are additional advantages of intervening before behaviours become personality traits, and there is a strong element of sensitivity shown in constructing the intervention to reduce the stigma of attending sessions. The balance of these factors is persuasive in accepting that ethical issues of gaining fully informed consent for at-risk interventions are satisfied.

Ethical Issues in Consent:

It is recognised that participants should give informed consent before they take part in a study and this should involve a formal stage. This study did not use a written consent from participants, partly because those involved were minors, but more because of the methodological sensitivities in running a secondary prevention intervention and the reluctance to label children as disaffected before their symptoms have reached a certain level of severity. Research with children requires their consent and that of their parents, and investigators must ensure participants are aware of their right to withdraw from the study even after the data have been collected. Withholding information or misleading participants is unacceptable if they are likely to show unease or object once debriefed.

In running the AGI intervention and research programmes fully informed consent was sought from all parties. Once identified, consent for children's participation in the intervention and the research was sought using the school systems of parental contact. Using the school context to conduct a programme of intervention with the aim of providing remediation for problematic behaviours, but also to prevent such behaviours registering on official codes of concern required sensitivity. Parents of some children were only too aware that the behaviour of their child was a problem at school, some parents however, especially those of children held back by their lack of confidence and inability to assert themselves in an active classroom, would be unaware that any problem necessitating intervention existed. Furthermore, the latter group of parents would naturally be disconcerted by the inclusion of their child in a group for naughty or disaffected children. In addition, the criteria are deliberately phrased to be able to include

brighter children whose social skills were lacking and whose peer relationships provided a source of concern. These considerations make it imperative for the success of this secondary prevention project that fully informed consent is obtained with sophistication and sensitivity. The consent procedure was structured to include three stages:

1. A letter from the head teacher to individual parents
2. A meeting in school for parents to learn more about the project
3. Individual meetings with potential child participants prior to the start of interventions.

The letter from the head teacher was simply worded and emphasised the positive aspects of involvement in the project, saying that it was an addition to the regular curriculum and the intervention itself contained much that could be considered as extra-curricula arts provision. The intervention was characterised as providing a focus for group skills, reducing conflict in the class and releasing the potential of students involved. The letter concluded by offering two dates where parents could attend an after-school meeting to learn more about the project, and included a telephone number of the AGI project leader for individual concerns. It was explained that the children would be allocated to one of two groups and the content of the groups would be different, but the hope was that the outcomes would be similar.

The meetings in school were informal and included a short presentation of the intervention aims and the research process. The aims of the research were spelt out and the positive aspects of involvement for each individual child were emphasised. Care was given to underline the truth of the intention not to generate a group that could be characterised as a group for the naughty children or those with anger management problems. No negative stigma could be attached to the group and it was said to the parents that running these groups in school had rarely produced anything but enthusiasm and willingness to attend. It was again explained that the children would be allocated to one of two groups and the content of the groups would be different, but the hope was that the outcomes would be similar.

Individual meetings with children were planned prior to the beginning of any assessment or intervention started. These sessions were attended by children in pairs to reduce the intimidation of having to speak to a stranger to school. During the meetings the consent of the child was sought for their involvement in a group intervention and the content of

the group that they had been allocated to was explained. The conversations were conducted with the intent to reduce levels of anxiety prior to the interventions starting and to generate a level of anticipation about being involved with the programme. At the start of the interventions the participants agree and sign a contract governing content and conduct relating to the group. This is a process that takes some time in the group and is thoughtfully completed by children and facilitator. In addition during the early sessions of the group, a question will be asked (if not from the children, then from the facilitator) about why the children are all present, what have they done to make their presence necessary? In this open airing of the inclusion criteria children quickly become aware that the group cannot be easily categorised as a bunch of 'naughty children' (or similar) and become more comfortable with their membership of the group. These 'soft' processes alongside the more formal consent procedures help in ensuring that consent procedures are adequate for child and parent.

Section 4.5: Issues of Reflexivity and Bias

Section 4.1 described the research design in some detail, looking at the framework and concerns that shaped the research decisions. As part of that exposition was a section about researcher values, with the recognition that it is not possible to keep the researcher values totally in check. Researcher values can be seen to potentially intrude in choice of research area, formulation of research question, choice of method, formulation of research design and data collection techniques, implementation of data collection, analysis of data, interpretation of data and of course the conclusions - in all aspects of a study like this, in other words. This subjects all aspects of the study to bias, but most tellingly its outcomes and implications for practice and theory.

On the face of it, a quantitative study like the AGI study should stand a better chance of being value-free and objective. Certainly the quantitative tradition would argue that research that simply reflected the personal biases of its practitioners could not be considered valid and scientific because it was bound up with the subjectivities of its practitioners. Another, more credible position in relation to the whole question of values and bias is to recognise and acknowledge that research, whether quantitative or qualitative, cannot be value-free, and the task is to ensure that there is no untrammelled incursion of values in the research process. This quality of self-reflection is increasingly considered vital to quality research processes and the exhibition of reflexivity by the

researchers about the part played by values and personal biases is increasingly essential.

Reflexivity has several meanings in the social sciences, but here is taken to mean that social researchers should be reflective about the implications of their methods, values biases and decisions for the knowledge of the social world they generate. Reflexivity should generate something of the nuts and bolts of research process as distinct from the often sanitized portrayal in research articles. In an extension of reflexivity in this sense is the acknowledgement of the role of the researcher as part and parcel of the construction of knowledge. In this sense, there is scepticism about the notion that the researcher is someone who extracts knowledge from observations and transmits the knowledge to an audience. Rather, the researcher is viewed as implicated in the construction of knowledge through the stance that he or she assumes in relation to the observed and through the ways in which an account is transmitted in the form of a text. This understanding entails an acknowledgement of the significance and implications of the researcher's choices as both observer and writer.

The researcher's awareness of his own bias is expressed through the discussion of research design and the inversion adoption of a 'bottom-up framework' to guide methodological choices. In that section some time is spent explaining to the reader about the practical considerations that shaped the research choices before going on to a detailed discussion of the values of the researcher and the potential implications this has on the areas of reporting outlined at the beginning of this section. In addition each chapter is begun with an introduction to contextualise the chapter and to provide it with some narrative. In this Introduction the fictionalised scenarios that begin Chapter One are kept to the forefront, an attempt to locate the chapter content within the context of so-called problem children. The aim is to give a clear expression to the sympathy that the researcher feels with his subject group and the biases he is subject to when engaging with the material of the chapter. This introduction together with the reflective commentary that concludes each chapter should contextualise this sympathy together with the other biases in the discourse of reflexivity that is outlined above.

Reflective Commentary

This methodology section as a whole conveys a great deal of the complexity and organisational scale of the AGI study. It is difficult to ascribe a narrative to the collection of detailed explanations that form the chapter, but if it were to tell a story, it would be one of determination to provide a premium research process on a budget reality. The gap between the funds and expertise has been filled with hard work and tenacity. The attempt at providing this premium process can be seen to be largely successful, but is compromised in key areas. It seems clear that an experimental programme was the aim, with an investment in quantitative data and subsequent analysis. The structure of the subject group and concerns for ecological validity meant that this standard was not completely met, and although randomisation and control procedures proceeded, the methodology does fall short of strict experimental conditions. While it may be questioned whether the desire to fit the subjects and nature of the AGI study into the experimental paradigm was an unrealistic project from the start, it seems clear that the intention to provide data with maximum persuasive impact for the research community was the driver and this position cannot be decried. One wonders if a commensurate investment had been made into establishing qualitative methods of investigating causality and process as the *primum mobile* of the study, the result would have been a better fit with the intervention and youth community, but it cannot be denied that the intention to inform change by aspiring to gold-standards of research methodology (however tainted) is a worthy one.

The chapter began with a statement to the effect that the detail of the methodological choices made would include not only the notion of choice between alternatives, but also the principles of rigour that underpinned those choices. In this way bias is acknowledged, and navigation through researcher preferences is made possible. It is mentioned several times that the methodological choices help solve the problems of how to generate persuasive data within the constraints of time resources and expertise available. The resulting data will have limitations, not just those of reliability extensively considered in the chapter, but also in terms of the lack of phenomenological account of change in the subject group. Attempts to supplement questionnaire data to redress the balance are clearly set out and the resulting methodology, while being something of an odd hybrid mix of experimental design and field study, certainly allows an impressive volume of information to be collated in a design of commendable rigour.

Persuading the schools, children and parents of the inner-city estates where the study was active presented a number of challenges. The schools themselves were wary, but all were convinced of the need to tackle disaffection in their classrooms and of the good intentions of the project. That is to say, it was never perceived that the study was hit-and-run research, but a serious attempt to improve the lot of children and teachers. It remains difficult to establish how much of the full research aims were effectively communicated to parents, and it is equally clear that they were the respondent group with the least patience with the process. On one occasion, a research assistant (RA) was out collecting data on the Meadow Well Estate in North Shields. After trying a few doors with no success the (female) RA parked her hatchback outside a house where she noticed a fair bit of activity further down the street. The parent was in, but after a brief discussion, was found to be unwilling to complete the CBCL form. On her return to the car the RA found that the young people playing in the road had lit wood and paper under it. This fire was just catching on as she hurried to drive away from a developing and dangerous situation. As she did so the youths threw the lighted branches after her. The RA was surprisingly relaxed as she retold the tale, putting the episode down to experience. 'Every day is a school day' she was heard to remark. One might wonder if the schoolteachers of the young people involved would say the same thing...

Chapter Five: Results

Introduction

This chapter presents the data gathered through the methodology detailed in the previous chapter. Due to reasons explained above, the bulk of the data are quantitative and the presentation that follows is necessarily number-based as are the analyses. It is important to remember that this is a process of examination of the effectiveness of three intervention conditions, a no-intervention waiting-list condition, and the AGI and CSG interventions. The object of the exercise is to establish some kind of organising principles in the deployment of resources in meeting the challenge of disaffection. Each number relates ultimately to a child in a socially deprived area of the North East, and without becoming sentimental, it is a timely reminder as the involved process of establishing causation is engaged.

Data were gathered at 4 time points:

T1 – three months prior to intervention

T2 – immediately prior to intervention

T3 – immediately following the cessation of intervention

T4 – one year following the cessation of intervention.

Hypotheses 1-4

The hypotheses use the questionnaire data to examine the extent and nature of change in the following ways: considering both groups together (AGI+CSG), a non-intervention waiting-list period (T1-T2) is compared to the intervention period (T2-T3) in Hyp1. Still considering both groups together, change over the intervention period is compared to a follow-up period of 1 year (T3-T4) in Hyp2. Hyp3 compares the gains of each intervention group (AGI vs CSG) over the intervention period (T2-T3), while Hyp4 compares the progress of each intervention group over the follow-up period.

These hypotheses use two sets of questionnaires: The Teacher Report Form (TRF), Youth Self-Report (YSR) and Child Behaviour Checklist (CBC). These are widely used questionnaires in child and youth mental health clinical practice and research (Achenbach 1991). Scores are generated on a number of factors covering two broad

areas of Externalising (e.g. aggressive, delinquent behaviour referred to in the text as behavioural problems) and Internalising behaviours (e.g. anxiety and depression referred to in the text as emotional problems). Raw scores are usually transformed into 't'-scores, approximately based on normalised scores: a T-score of 50 represents the mean score for a population, 60 (1 standard deviation above the mean) represents the 'borderline clinical' cut-point.

The Multidimensional Self Concept Scale (MSCS) is a self-rating scale consisting of 150 items, grouped into 6 sub-scales of high reliability (Cronbach's alpha for the total scale 0.98, and 0.85 or above for each subscale). Test-retest reliability is 0.90. The MSCS taps self-perceptions relating to important areas of school and family life (Bracken 1992). The scores on the questionnaire are for positive self-esteem, with greater scores indicating improvement, however, in order to give consistency in reporting data from the MSCS and the Achenbach problem scales, the scores have been inverted so that a declining mean score in the following report indicates improvement.

Section 5.1: Hypothesis 1

Question 1. Does working in a group affect a child's emotional/behavioural profile?
Do children gain in the skills reported by questionnaires when they are involved in group interventions?

Question 1(sub). Are any gains in emotional/behavioural profile reported during the group intervention greater than those over a similar period when no intervention is taking place?

If you compare the time when children have been working in a group with the time when they have not, will you notice a difference in their behavioural and emotional ratings?

What is the effect of working in a small group compared to usual school class experience? To answer this, both the AGI and the CSG groups are considered together, and time points T1, T2 and T3 are compared. It is predicted that positive feedback, group purpose and forces of group identity will have an effect of reducing problem behaviours and encouraging self-esteem. Hypothesis No.1 therefore states that it is expected that these interventions will have positive effects over whole period (T1 – T3). As this period contains a waiting-list period (T1 – T2) during which there was no

intervention, the sub hypothesis is that the gains during this T1 – T2 period will be less than the T2 – T3 intervention period.

Hypothesis 1: According to parents, teachers and self-report, working in groups reduces behavioural and emotional problems and promotes self-esteem as reported from initial assessment to post-intervention (T1 – T3).

The Hypothesis is split into four parts (a – d) each relating to a different dataset.

- a) According to child self-report (YSR), working in groups reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).
- b) According to child self-report (MSCS), working in groups regard promotes self-esteem as reported from initial assessment to post-intervention (T1 – T3).
- c) According to teacher report (TRF), working in groups reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).
- d) According to parent report (CBC), working in groups reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).

Hypothesis 1 (sub): According to parents, teachers and self-report, the gains over the no-intervention period (T1 – T2) will be less than over the subsequent intervention period (T2 – T3).

The Hypothesis is split into four parts (a – d) each relating to a different dataset.

- a) According to child self-report (YSR), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).
- b) According to child self-report (MSCS), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).
- c) According to teacher report (TRF), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).

- d) According to parent report (CBC), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).

In the detail that follows, each part is expressed in its positive and null form.

Mean scores are reported at three time points (T1: baseline, T2: pre-intervention and T3: post-intervention). ANOVA scores and Effect sizes are calculated for the difference in scores for three time periods, T1-T3: the period of waiting until the end of intervention, T1-T2: the period of the waiting list control and T2-T3: the period of intervention.

Hypothesis 1a & 1(sub)a

Under this hypothesis we report on the findings from the Youth Self Report (Achenbach 1991) a problem scale completed by the children who received the groupwork interventions.

Experimental Hypothesis 1a:

“According to child self-report (YSR), working in groups reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).”

Null Hypothesis 1a:

“According to child self-report (YSR), working in groups produces no reduction in behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).”

Experimental Hypothesis 1(sub)a:

“According to child self-report (YSR), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).”

Null Hypothesis 1(sub)a:

“According to child self-report (YSR), gains over the intervention period (T2 – T3) will be equal to or less than those made in the no-intervention period (T1 – T2).”

Findings for Hypothesis 1a and 1(sub)a:

Table 1a Descriptives

<u>YSR</u>	T	Mean	SD	N	df
Total T-Score	1	60.39	9.19	44	2
	2	54.82	12.41	44	
	3	52.36	10.46	44	
Internal T- Score	1	61.32	8.57	44	2
	2	56.18	9.94	44	
	3	52.57	9.35	44	
External T-Score	1	56.75	10.13	44	2
	2	52.80	13.17	44	
	3	52.57	11.3	44	

Table 1a Analysis

YSR	Repeated Measures ANOVA & Effect Size (ES)								
	T1-T3			T1-T2			T2-T3		
	F	p<	ES	F	p<	ES	F	p<	ES
Total T-Score	11.36	0.01 (0.000)	0.82	9.4	0.01 (0.004)	0.52	2.05	ns (0.159)	0.22
Internal T-Score	17.22	0.01 (0.000)	0.98	10.76	0.01 (0.002)	0.56	6.14	0.05 (0.017)	0.38
External T-Score	3.88	0.05 (0.024)	0.39	4.97	0.05 (0.031)	0.34	0.02	ns (0.894)	0.02

Reporting Results for Hypothesis 1a and 1(sub)a:

The population of drama and curriculum groups is considered here as a whole.

Hypothesis 1a is concerned with the overall time period from the beginning of the waiting list period T1 to the end of the intervention period T3, while 1(sub)a is concerned with comparing T1-T2 with T2-T3

For the full period: T1-T3

Total problem T-scores: the total mean scores drop sharply, indicating a substantial improvement in self-reported scores taking place during this time T1 – T3. The p-value

of 0.000 shows that this improvement is highly statistically significant, and the effect size is moderate.

Internalising T-scores: the Internalising mean scores also drop sharply, indicating that children feel that their emotional problems improving. This is highly statistically significant (0.000) and the effect size of 0.98 is moderate.

Externalising T-scores: these T3 scores are of the same order as the Internalising scores but are not reported to be as severe at T1. Consequently the p-value reveals less statistical significance (0.024), however $p < 0.05$. The effect size of 0.39 is small.

Going on to look at whether this improvement is happening in the period T1-T2 or the intervention period T2-T3 is the central question that the Hypothesis 1(sub)a seeks to answer.

For the waiting-list period: T1-T2

Total problem T-Scores: the scores drop from 60 to 54, a substantial improvement. The statistical significance level is high with a p-value of 0.004 and the effect size at 0.52 is small.

Internalising T-scores: the mean scores drop from 61 to 56, a highly statistically significant change ($p=0.002$) and the magnitude of change is small at 0.56.

Externalising T-scores: the mean scores drop from 56 to 52, not such a highly statistically significant change ($p=0.031$) but $p < 0.05$. The effect size at 0.34 is small.

For the intervention period: T2-T3

Total problem T-scores: the scores drop from 54 to 52 not a statistically significant change ($p=0.159$) and a small magnitude of change $ES=0.22$

Internalising T-scores: The mean scores drop from 56 to 52, $p < 0.05$ ($p=0.017$) with an effect size small at 0.38.

Externalising T-scores: the mean scores fall only marginally from 52.80 to 52.57, not significant ($p=0.894$) and of a trivial magnitude $ES=0.02$.

Summary:

These data give compelling evidence to accept that self-reported Achenbach change in the waiting-list no-intervention period is more likely to happen for more children in greater amounts than self-reported change over the intervention period. This holds true

for total scores and Externalising scores, but there are indications that the self-reported emotional problems are improving over the intervention period.

On the basis of this evidence the Null Hypothesis 1a can be rejected and the Experimental Hypothesis supported.

“According to child self-report (YSR), working in groups reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).”

However, the experimental hypothesis 1(sub)a is rejected and the null hypothesis 1(sub)a is accepted:

“According to child self-report (YSR), gains over the intervention period (T2 – T3) will be equal to or less than those made in the no-intervention period (T1 – T2).”

Interpretation of Data and Discussion of Hypothesis 1a and 1(sub)a:

These data give little room for alternative interpretations. From the children's point of view, their Externalising aggressive problems and their Internalising emotional problems get better when there is no intervention going on. When involved in groupwork, the gains in Externalising problems are sustained, and Internalising problems continue to improve with high statistical significance, and of a small magnitude $ES = 0.39$.

This indication of the effects of groupwork on the problems reported by children is important, but its importance is overshadowed by the gains that are of greater magnitude for the waiting list period. This fact is puzzling and counter-intuitive.

In searching for explanations without reference to other findings, at this stage it is possible to speculate that an unfamiliar process might 'surprise' the children into providing honest high scores in the first instance T1. These scores might then be moderated or toned down at T2 by the knowledge that they are about to take part in a positive intervention programme of some sort. It is reasonable to assume that there will be a degree of expectation about the term's work and a feeling of being special, singled out for positive attention that will be active in their scorings of their own problems at T2. How they feel about themselves is an important factor in the reporting of problems, perhaps more in self-reports than in other reports. There is a reasonable possibility that

the scores at T2 register artificially positive behaviour, but those at T1 and T3 are a true reflection of their problems.

On the other hand it is possible to argue for an effect of selection operating in opposite directions: it could be that children feel it necessary to suppress their initial scores, keeping them artificially low so as to avoid being thought of as problematic: something that children might do if they were keen to participate in the group sessions. Certainly the positive presentation of the interventions could encourage enthusiasm, but it is difficult to know what children were privately thinking.

Attrition:

The numbers of children involved is 44 present at all three assessment points from a population of 132. The difference is accounted for by a factor in experimental design that meant that the assessment programme started with interventions and waiting list assessments together, meaning that the first cohort of children did not have the chance to be included in the waiting list condition.

Hypothesis 1b & 1(sub)b

This relates to the supplementary self-assessment measure, the Multi-Dimensional self-concept Scale (MSCS). This measure is not problem orientated like the Achenbach and gives scores of positive and pro-social measures. This provides an important check to the emphasis on problem questions being asked of children who may well then begin to wonder unnecessarily about behaviours. It also provides the possibility of scores for items that will not register on the Achenbach. A child may score the Achenbach as all zeros, leaving no room for improvement, but it is difficult to leave the MSCS with a similar lack of data. MSCS scores are inverted to provide consistency with the Achenbach scales, therefore declining scores indicate improvement.

Experimental Hypothesis 1b:

“According to child self-report (MSCS), working in a small group promotes self-esteem as reported from initial assessment to post-intervention (T1 – T3).”

Null Hypothesis 1b:

“According to child self-report (MSCS), working in a small group results in no promotion self-esteem as reported from initial assessment to post-intervention (T1 – T3).”

Experimental Hypothesis 1(sub)b

“According to child self-report (MSCS), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).”

Null Hypothesis 1(sub)b:

“According to child self-report (MSCS), gains over the intervention period (T2 – T3) will be equal to or less than those made in the no-intervention period (T1 – T2).”

Findings for Hypothesis 1b and 1(sub)b:

Table 1b Descriptives

	t	Mean	SD	N	df
Total T-Score	1	55.33	7.14	27	2
	2	50.67	8.75	27	
	3	50.19	8.60	27	
Academic T- Score	1	53.00	7.51	28	2
	2	49.5	9.96	28	
	3	49.64	9.39	28	
Affect T-Score	1	54.75	8.77	28	2
	2	49.82	8.97	28	
	3	50.07	7.73	28	
Competence T-Score	1	57.64	8.28	28	2
	2	53.32	8.23	28	
	3	51.79	8.33	28	
Family T-Score	1	53.32	6.53	28	2
	2	49.29	9.53	28	
	3	50.14	8.96	28	
Physical T- Score	1	53.04	6.81	28	2
	2	50.86	7.12	28	
	3	48.67	8.05	28	
Social T-Score	1	55.29	8.45	28	2
	2	49.96	9.86	28	
	3	49.29	7.88	28	

Table 1b Analysis

	Repeated Measures ANOVA & Effect Size (ES)								
	T1-T3			T1-T2			T2-T3		
	F	p<	ES	F	p<	ES	F	p<	ES
Total T-Score	8.147	0.01 (0.001)	0.66	9.86	0.01 (0.004)	0.59	0.122	ns (0.730)	0.06
Academic T-Score	3.72	0.05 (0.031)	0.40	5.40	0.05 (0.028)	0.41	0.010	ns (0.920)	-0.02
Affect T-Score	4.58	0.05 (0.015)	0.57	6.05	0.05 (0.021)	0.56	0.034	ns (0.856)	-0.03
Compet-ence T-Score	8.62	0.01 (0.001)	0.71	6.975	0.05 (0.014)	0.52	1.59	ns (0.219)	0.19
Family T-Score	4.26	0.05 (0.019)	0.42	8.07	0.01 (0.008)	0.51	0.267	ns 0.609	-0.09
Physical T-Score	4.55	0.05 (0.015)	0.59	2.02	ns (0.167)	0.31	2.22	ns (0.148)	0.29
Social T-Score	7.03	0.01 (0.002)	0.74	13.21	0.01 (0.001)	0.58	0.127	ns (0.724)	0.08

Reporting Results for Hypothesis 1b & 1(sub)b:

The population of drama and curriculum groups are considered here as a whole. Hypothesis 1b is concerned with the overall time period from the beginning of the waiting list period T1 to the end of the intervention period T3, while 1(sub)b is concerned with comparing T1 – T2 with T2 – T3.

Total Period T1-T3

Total MSCS T-scores
The mean scores over this period fall from 55.33 to 50.19 and with a p-value of 0.001 is highly statistically significant, well below the p<0.01 level of probability. The magnitude of change (effect size) is moderate at 0.66: over half a standard deviation. This indicates a magnitude of change that is very unlikely to have happened by chance.

Academic Dimension T-Scores

The perception the children have of their academic performance shows an improvement, with mean scores falling from 53 to 49.64. This change is statistically significant at the 0.05 level ($p=0.031$) and a small effect size of 0.40 of a standard deviation.

Affect Dimension T-scores

The feeling dimension shows an improvement, with mean scores falling from 54.75 to 50.07. This change is statistically significant at the $p<0.05$ level with a small effect-size of 0.57 of a standard deviation

Competence Dimension T-scores

Children rating their competence over the period show a fall in mean scores from 57.64 to 51.79, a change statistically significant at the $p<0.01$ level. The moderate effect size is 0.71 of a standard deviation

Family Dimension T-scores

Children's perception of their family situation improves, with mean scores falling from 53.32 to 50.14, a change statistically significant at the $p<0.05$ level. The small effect size is 0.42 of a standard deviation.

Physical Dimension T-scores

The children's rating of their physical appearance and ability improves, with mean scores falling from 53.04 to 48.67, a change statistically significant at the $p<0.05$ level. The small effect size is 0.59 of a standard deviation.

Social Dimension T-scores

The children's perception of their social skills also shows an improvement with mean scores falling from 55.29 to 49.29, a change statistically significant at the $p<0.01$ level. The moderate effect size is 0.74 of a standard deviation.

For Waiting-List Period T1 - T2

Total MSCS T-scores

Total mean scores over the waiting list period fall from 55.33 to 50.67, a change statistically significant at the $p<0.01$ level. The moderate effect size is 0.66 of a standard deviation.

Academic Dimension T-Scores

There is a statistically significant improvement over the waiting-list period on this measure ($p=0.028$), mean scores falling T1-T2: 53.00 - 49.50. The small effect size is 0.41 of a standard deviation.

Affect Dimension T-scores

There is a statistically significant improvement over the waiting-list period on this measure ($p=0.021$), mean scores falling T1-T2: 54.75- 49.82. The small effect size is 0.56 of a standard deviation.

Competence Dimension T-scores

There is a statistically significant improvement over the waiting-list period on this measure ($p=0.014$), mean scores falling T1-T2: 57.64 – 53.32. The small effect size is 0.52 of a standard deviation.

Family Dimension T-scores

There is a statistically significant improvement over the waiting-list period on this measure ($p=0.008$), mean scores falling T1-T2: 53.32 - 49.29. The small effect size is 0.51 of a standard deviation.

Physical Dimension T-scores

There is no statistically significant improvement over the waiting-list period on this measure ($p=0.167$), mean scores falling T1-T2: 53.04 – 50.86. The small effect size is 0.31 of a standard deviation.

Social Dimension T-scores

There is a statistically significant improvement over the waiting-list period on this measure ($p=0.001$), mean scores falling T1-T2: 55.29 - 49.96. The small effect size is 0.58 of a standard deviation.

For Intervention Period T2 – T3

Total Problem T-scores

There is no statistically significant change over the intervention period on this measure ($p=0.73$), mean scores hardly change: T2-T3: 50.67 – 50.19. The trivial effect size is 0.06 of a standard deviation.

Academic Dimension T-Scores

There is no statistically significant change over the intervention period on this measure ($p=0.92$), mean scores hardly change: T2-T3: 49.50 – 49.64. The effect size is trivial.

Affect Dimension T-scores

There is no statistically significant change over the intervention period on this measure ($p=0.86$), mean scores hardly change: T2-T3: 49.82 – 50.07. The effect size is trivial.

Competence Dimension T-scores

There is no statistically significant change over the intervention period on this measure ($p=0.22$), mean scores hardly change: T2-T3: 53.32 – 51.79. The effect size is trivial.

Family Dimension T-scores

There is no statistically significant change over the intervention period on this measure ($p=0.61$), mean scores hardly change: T2-T3: 49.29 – 50.14. The effect size is trivial.

Physical Dimension T-scores

There is no statistically significant change over the intervention period on this measure ($p=0.15$), mean scores hardly change: T2-T3: 50.86 – 48.67. The effect size is trivial.

Social Dimension T-scores

There is no statistically significant change over the intervention period on this measure ($p=0.72$), mean scores hardly change: T2-T3: 49.96 – 49.29. The effect size is trivial.

Summary:

There is statistically significant change on MSCS total scale scores and all sub-scale scores over the period of initial assessment to post-intervention (T1 – T3) with effect size scores ranging from small to moderate. It is possible to reject the null hypothesis 1b.

Experimental Hypothesis 1b:

“According to child self-report (MSCS), working in a small group promotes self-esteem as reported from initial assessment to post-intervention (T1 – T3).”

There is no statistically significant change over the intervention period T2 – T3 on any of the scales reported by the MSCS questionnaire. Significant change is present on all scales, with the exception of the physical scale in the waiting list period T1 – T2. On this evidence the experimental hypothesis is rejected.

Null hypothesis 1(sub)b:

“According to child self-report (MSCS), gains over the intervention period (T2 – T3) will be equal to or less than those made in the no-intervention period (T1 – T2).”

Interpretation of Data and Discussion of Hypothesis 1b & 1(sub)b:

From the children's own perception, their self-esteem improves over the no-intervention waiting-list period and remains constant without falling significantly over the intervention period. These findings like those of the other self-report, the YSR, are counter-intuitive and some of the same explanations of expectation effects are relevant here also. There does seem to be a significant effect that arises out of having been selected for special treatment. Without population data from the rest of the class to compare it to it is difficult to say whether this effect is part of the effect of class environment. The data comes from a range of schools at different times of the year, so it is difficult to attribute it to time of year effects or to ethos effects particular to one school.

The gains made in self-esteem are sustained over the intervention period, in other words it is possible to present evidence that the children did not feel let down by the interventions.

Multiple Testing:

Using logarithms developed by Sakoda, Cohen and Beal it is possible to predict the probability of obtaining n or more significant statistics out of N calculated statistics (Sakoda, Cohen, & Beal 1954). In this case, the number of calculated statistics is the number of scores i.e. six (the MSCS total score is a composite of the scale scores and not therefore an independent variable). These are tested over three time periods giving a total of 18 calculated statistics. Of these statistics calculated for Hypothesis 1b and 1(sub)b 14 are significant, some at $p < .05$ some at $p < .01$. Using the developed logarithms, the chance probability of obtaining this number of significant statistics is well under .001.

Attrition:

The numbers are lower for this hypothesis than for the others, because the first cohort of children did not receive the waiting list questionnaire, as the intervention programme

begun at the same time as the assessment programme. With smaller figures like this, the significance levels achieved are all the more compelling.

Hypothesis 1c & 1(sub)c

The Teacher Report Form is part of the Achenbach group of questionnaires and reports problem scales in the same way as the YSR. The interventions are school-based and the teachers involved in the selection of children, so it is reasonable to expect that teachers will register most change.

Experimental Hypothesis 1c:

"According to teachers, working in a small group reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1-T3)."

Null Hypothesis 1c:

"According to teachers, working in a small group produces no reduction in behavioural and emotional problems as reported from initial assessment to post-intervention."

Experimental Hypothesis 1(sub)c:

According to teacher report (TRF), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2)."

Null Hypothesis 1(sub)c:

According to teacher report (TRF), gains over the intervention period (T2 – T3) will be equal to or less than those made in the no-intervention period (T1 – T2)."

Findings for Hypothesis 1c &1(sub)c:

Table 1c Descriptives

TRF	t	Mean	SD	N	do
Total T-Score	1	58.50	10.64	82	2
	2	57.90	11.40	82	
	3	52.32	10.71	82	
Internal T- Score	1	58.99	11.92	82	2
	2	57.94	12.64	82	
	3	52.35	12.41	82	
External T-Score	1	57.30	12.00	81	2
	2	57.40	11.32	81	
	3	52.77	11.29	81	

Table 1c Analysis

TRF	Repeated Measures ANOVA & Effect Size (ES)								
	T1-T3			T1-T2			T2-T3		
	F	p<	ES	F	p<	ES	F	p<	ES
Total T-Score	33.50	0.01 (0.000)	0.58	0.75	ns (0.390)	0.05	38.79	0.01 (0.000)	0.5
Internal T- Score	28.57	0.01 (0.000)	0.55	1.53	ns (0.219)	0.09	28.66	0.01 (0.000)	0.45
External T-Score	13.69	0.01 (0.000)	0.39	0.28	ns (0.867)	-0.01	15.98	0.01 (0.000)	0.41

Reporting Results for Hypothesis 1c:

The population of drama and curriculum groups are considered here as a whole.

Hypothesis 1c is concerned with the overall time period from the beginning of the waiting list period (T1) to the end of the intervention period (T3), while 1(sub)c is concerned with comparing T1-T2 with T2-T3.

For Full Period T1-T3

Total Problem T-scores

There is a statistically significant improvement over the full period on this measure (p=0.000), mean scores falling T1-T3: 58.50 – 52.32. The small effect size is 0.58 of a standard deviation.

Internalising Problem T-Scores

There is a statistically significant improvement over the full period on this measure ($p=0.000$), mean scores falling T1-T3: 58.99 – 52.35. The small effect size is 0.54 of a standard deviation.

Externalising Problem T-scores

There is a statistically significant improvement over the full period on this measure ($p=0.000$), mean scores falling T1-T3: 57.30 – 52.77. The small effect size is 0.39 of a standard deviation.

For Waiting-List Period T1-T2

Total Problem T-scores

There is no statistically significant improvement over the waiting-list period on this measure ($p=0.390$), mean scores showing only slight variability T1-T2: 58.50 – 57.90. The effect size is trivial.

Internalising Problem T-Scores

There is no statistically significant improvement over the waiting-list period on this measure ($p=0.219$), mean scores showing only slight variability T1-T2: 58.99 – 57.94. The effect size is trivial.

Externalising Problem T-scores

There is no statistically significant improvement over the waiting-list period on this measure ($p=0.867$), mean scores showing only slight variability T1-T2: 57.30 – 57.40. The effect size is trivial.

For Intervention Period T2-T3

Total Problem T-scores

There is highly statistically significant improvement over the intervention period on this measure ($p=0.000$), mean scores falling T2-T3: 57.90 – 52.32. The small effect size is 0.51 of a standard deviation.

Internalising Problem T-Scores

There is highly statistically significant improvement over the intervention period on this measure ($p=0.000$), mean scores falling T2-T3: 57.94 – 52.35. The small effect size is 0.45 of a standard deviation.

Externalising Problem T-scores

There is highly statistically significant improvement over the intervention period on this measure ($p=0.000$), mean scores falling T2-T3: 57.40 – 52.77. The small effect size is 0.41 of a standard deviation.

Summary:

The problem scores reduce over the full period T1-T3, the change being highly significant on Total, Internalising and Externalising scales. The effect sizes of these reductions are small for Total, Externalising and Internalising scores but above or close to the 0.4 average for educational interventions. This gives us evidence to reject the Null Hypothesis 1c.

Experimental Hypothesis 1c:

“According to teachers, working in a small group reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1-T3).”

The significant change occurs over the intervention period T2-T3 of small magnitude but above the 0.4 average for educational Interventions, with no significant change reported on any scale over the waiting list period T1-T2. On this evidence it is possible to reject the Null Hypothesis 1(sub)c

Experimental Hypothesis 1(sub)c:

“According to teacher report (TRF), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).”

Interpretation of Data and Discussion of Hypothesis 1c and 1(sub)c:

The data from the teachers confirm the expectations of the research team, that there is change associated with Intervention, not duplicated in the waiting-list no-intervention period. The teacher data do not follow the data from the children. The extent of the change is of the order of half a standard deviation on each scale (Total problem, Internalising and Externalising problems) over the intervention period T2 – T3 with a very high level of statistical significance ($p=0.000$). This change is compared to the waiting list period where there is no significant change and the effect sizes are trivial.

The teachers clearly see improvement associated with a type of groupwork that is not present in the usual classroom situation.

At these time points the teachers completing each form will have been the same. There is a potential bias in the reported data due to the need to get the permission of teachers by informing them of the aims and methodology of the research process. Teachers will have been aware that the waiting list period served as a no-intervention control and it is possible to speculate that they were subject to bias based on the perception that to report the effectiveness of groupwork will have increased their chances of getting longer term assistance with groups of at-risk children, inevitably those who create extra work for teachers. Teachers are of course human, and are not immune from thinking in the way proposed above. As well as being human, however, teachers are professionals and it is hoped that professional standards acted to redress any possible bias.

Attrition:

The numbers of 81 and 82 respondents for this questionnaire are higher than those for the YSR because the questionnaires are easier to administer, teachers are more reliable in completing questionnaires, one teacher completes a form for several children and some cohorts had only TRF assessments made.

Hypothesis 1d & 1(sub)d:

Under this hypothesis we report findings of the Child Behaviour Checklist completed by parents. This is an Achenbach problem scale.

Experimental Hypothesis 1d:

“According to parent report (CBC), working in a small group reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).”

Null Hypothesis 1d:

“According to parent report (CBC), working in a small group produces no reduction in behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).”

Experimental Hypothesis 1(sub)d:

“According to parent report (CBC), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).”

Null Hypothesis 1(sub)d:

“According to parent report (CBC), gains over the intervention period (T2 – T3) will be equal to or less than those made in the no-intervention period (T1 – T2).”

Findings for Hypothesis 1d &1(sub)d

Table 1d Descriptives

CBCCL	t	Mean	SD	N	do
Total T-Score	1	55.84	10.73	31	2
	2	56.13	10.71	31	
	3	52.94	10.25	31	
Internal T- Score	1	54.42	10.56	31	2
	2	54.52	11.19	31	
	3	52.94	9.62	31	
External T-Score	1	55.35	11.62	31	2
	2	55.81	10.67	31	
	3	52.42	9.43	31	

Table 1d Analysis

	Repeated Measures ANOVA & Effect Size (ES)								
	T1-T3			T1-T2			T2-T3		
	F	p<	ES	F	p<	ES	F	p<	ES
Total T-Score	2.87	ns (0.064)	0.28	0.06	ns (0.808)	-0.03	3.94	ns (0.056)	0.31
Internal T-Score	0.42	ns (0.660)	0.15	0.003	ns (0.957)	-0.01	0.61	ns (0.442)	0.15
External T-Score	2.71	ns (0.075)	0.28	0.14	ns (0.709)	-0.04	4.71	0.05 (0.038)	0.34

Reporting Results for Hypothesis 1d and 1(sub)d:

The population of drama and curriculum groups are considered here as a whole. Hypothesis 1d is concerned with the overall time period from the beginning of the waiting list period T1 to the end of the intervention period T3, while 1(sub)d is concerned with comparing T1 – T2 with T2 – T3.

For Full Period T1-T3

Total Problem T-scores

There is no statistically significant improvement over the full period on this measure (p=0.064). Mean scores show variability T1-T3 falling from 55.84 – 52.94. The small effect size is 0.28 of a standard deviation.

Internalising Problem T-Scores

There is no statistically significant improvement over the full period on this measure (p=0.660). Mean scores show variability T1-T3 falling from 54.42 – 52.94. The effect size is trivial.

Externalising Problem T-scores

There is no statistically significant improvement over the full period on this measure (p=0.075). Mean scores show variability T1-T3 falling from 55.35 – 52.42. The effect size is small.

For Waiting-List Period T1-T2

Total Problem T-scores

There is no statistically significant improvement over the waiting-list period on this measure (p=0.81). Mean scores show slight variability T1-T2 rising from 55.84 – 56.13. The effect size is trivial.

Internalising Problem T-Scores

There is no statistically significant improvement over the waiting-list period on this measure (p=0.96). Mean scores show no variability T1-T2: 55.42 – 55.42. The effect size is trivial.

Externalising Problem T-scores

There is no statistically significant improvement over the waiting-list period on this measure ($p=0.71$). Mean scores show slight variability T1-T2 rising from 55.35 – 55.08. The effect size is trivial.

For Intervention Period T2-T3

Total Problem T-scores

There is statistically significant improvement just below $p<0.05$, over the intervention period on this measure ($p=0.06$), mean scores falling T2-T3: 56.13 – 52.94. The small effect size is 0.31 of a standard deviation.

Internalising Problem T-Scores

There is no statistically significant improvement over the intervention period on this measure ($p=0.442$), mean scores falling T2-T3: 54.52 – 52.94. The effect size is trivial.

Externalising Problem T-scores

There is statistically significant improvement over the intervention period on this measure ($p=0.038$), mean scores falling T2-T3: 55.81 – 52.42. The small effect size is 0.34 of a standard deviation.

Summary:

There is statistically significant change reported over the total period by the parents on Externalising scale, reflecting in the total scale, just outside the $p<0.05$ level ($p=0.064$ and 0.075 respectively). The effect sizes are of the order of 0.2 of a standard deviation and therefore show a small magnitude of change: this hypothesis is split, and on balance is accepted. Over the waiting list period T1-T2 the significance levels are much lower and effect sizes show behaviours getting worse, albeit marginally. Over the intervention period T2-T3 the Externalising scale reveals significant change at $p<0.05$ and the total scale is just outside the $p<0.05$ level (0.056). The effect size for both is small, but close to the 0.4 average for educational interventions, and the case for accepting this split finding is judged to be acceptable

Experimental Hypothesis 1d:

“According to parent report (CBC), working in a small group reduces behavioural and emotional problems as reported from initial assessment to post-intervention (T1 – T3).”

Experimental Hypothesis 1(sub)d:

“According to parent report (CBC), gains over the intervention period (T2 – T3) will be greater than those made in the no-intervention period (T1 – T2).”

Interpretation of Data and Discussion of Hypothesis 1D:

Parents do support the experimental hypothesis through the evidence provided by the effect size data, but significance is only reported in the case of Externalising problems, this occurring over the intervention period T2-T3. The feel of the data for the two time periods is very different, and judged to be experimentally significant: the parents register a change in their children's behaviour over the intervention period not mirrored in the waiting list period.

The importance of cross-contextual change in behaviour is a key concern of intervention and the evidence here to support these kind of improvements is suggestive, but nonetheless present. It is interesting that the parents report changes in Externalising problems first and foremost, when the children report no change in Externalising problems over the same period (YSR), but significant change in Internalising problems.

The children feel internal change, while the parents see external change. This appears to be evidence that the internal change felt by children translates into felt behavioural change. An alternative explanation might be that the parents have little access to or understanding of the child's feelings, leaving them unable to score internal states with any confidence.

The parents are also susceptible to bias, and it is reasonable to assume a level of anxiety in the initial stages of setting up the research programme, reflected perhaps in poor scores when no intervention was taking place. This could have been alleviated through contact with the researchers and when the children came home to report positively on the interventions. The more positive outlook might be reflected in more positive assessments of their child's behaviour at time point T3.

Attrition:

The numbers of 31 present at all three time points will have been affected by the cohort not participating in the waiting list condition because intervention and research started at the same time, also because the parent data are the most difficult to collect.

Section 5.2: Hypothesis 2

Question 2. Do gains in emotional/behavioural profile made while the groups are running last after the group is finished?

If you survey the respondents at intervals after the interventions are over will you still be able to detect the gains made?

Are the effects of working in a group sustained after intervention has finished? To answer this, both AGI and CSG groups are considered together, and time points T2, T3 and T4 at one-year follow-up are considered. The input common to both AGI and CSG groups is intended to have a lasting impact in the way a child behaves and based on the reading of the literature in Chapter One, this effect should be present one year after intervention. It is therefore predicted that any improvement for the intervention groups over the intervention period T2-T3 will be sustained one year later at T4. The hypothesis looks for significant improvement over intervention and no significant change over the follow-up period.

Hypothesis 2:

According to parents, teachers and self-report, gains made on emotional, behavioural and self-esteem measures over the intervention period (T2 – T3) will be sustained over the follow-up period (T3 – T4).

The hypothesis is split into four parts (a – d) each relating to a different instrument of measurement (TRF, CBC, YSR, MSCS).

- a) According to child self-report (YSR), gains made on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).**
- b) According to child self-report (MSCS), the gains on self-esteem measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).**
- c) According to teacher report (TRF), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).**

- d) According to parent report (CBC), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).

When examined below, each part is expressed in its positive and null form.

Mean scores are reported at three time points (T2: pre-intervention, T3: post-intervention and T4: 1-year post intervention). ANOVA scores and effect sizes are calculated for the difference in scores for three time periods, T2-T4: the period of intervention and follow-up, T2-T3: period of intervention and T3-T4 the period of follow-up one year post intervention. For the purposes of discriminating between the hypothesis and the null hypothesis it is the comparison of these latter two time periods that will be examined.

Hypothesis 2a

Hypothesis 2a examines the research question by looking at the responses of children filling in the Achenbach Youth Self Report at time points T2, T3 and T4.

Experimental Hypothesis 2a

“According to child self-report (YSR), gains made on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).”

Null Hypothesis 2a

“According to child self-report (YSR), gains made on emotional and behavioural measures over the intervention period (T2 – T3) are not sustained over the follow-up period (T3 – T4).”

Findings for Hypothesis 2a:

Table 2a Descriptives

YSR	t	Mean	SD	N	do
Total T-Score	2	54.27	12.25	77	2
	3	50.17	10.95	77	
	4	48.90	12.49	77	
Internal T- Score	2	54.70	11.81	77	2
	3	49.56	10.93	77	
	4	48.22	11.78	77	
External T-Score	2	53.04	11.81	77	2
	3	50.70	10.52	77	
	4	49.73	13.05	77	

Table 2a Analysis

YSR	ANOVA & Effect Size (ES)								
	T2-T4			T2-T3			T3-T4		
	F	p<	ES	F	p<	ES	F	p<	ES
Total T-Score	8.891	0.01 (0.001)	0.44	8.973	0.01 (0.004)	0.36	0.616	ns (0.435)	0.1
Internal T-Score	10.675	0.01 (0.000)	0.55	15.028	0.01 (0.000)	0.45	0.776	ns (0.381)	0.12
External T-Score	0.968	ns (0.055)	0.27	3.222	ns (0.077)	0.45	0.500	ns (0.482)	0.08

Reporting Results for Hypothesis 2a:
For Intervention Period T2-T3 & Follow-up period T3 – T4

Total Problem T-scores
There is statistically significant improvement over the intervention period on this measure (p=0.004), mean scores falling T2-T3: 54.27 – 50.17. The small effect size is 0.36 of a standard deviation.

There is no statistically significant improvement over the follow-up period on this measure (p=0.435), mean scores show little variability T3-T4, falling 50.17 – 48.90. The trivial effect size is 0.10 of a standard deviation.

Internalising Problem T-Scores

There is statistically significant improvement over the intervention period on this measure ($p=0.000$), mean scores falling T2-T3: 54.70 – 49.56. The small effect size is 0.45 of a standard deviation.

There is no statistically significant improvement over the follow-up period on this measure ($p=0.381$), mean scores show little variability T3-T4, falling 49.56 – 48.22. The trivial effect size is 0.12 of a standard deviation.

Externalising Problem T-scores

There is no statistically significant improvement over the intervention period on this measure ($p=0.077$), mean scores falling T2-T3: 53.04 – 50.70. The small effect size is 0.45 of a standard deviation.

There is no significant improvement over the follow-up period on this measure ($p=0.482$), mean scores show little variability T3-T4, falling 50.70 – 49.73. The trivial effect size is 0.08 of a standard deviation.

Summary:

There is no significant change on any measure of the YSR over the follow-up period T3-T4. Statistically significant improvement was seen on Internalising scores and Total scores over the intervention period T2-T3. The improvement in Externalising scores T2-T3 was judged to be marginally significant at $p=0.077$. The effect sizes were all positive for this period and on this evidence the null hypothesis is rejected

Experimental Hypothesis 2a:

“According to child self-report (YSR), gains made on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).

Interpretation of Data and Discussion of Hypothesis 2a:

Children clearly rate their own problems as improving over the intervention period and holding steady over the following year. This is an encouraging finding for the long-term effectiveness of small groupwork. The only equivocation that children express is in their assessment of their own Externalising behaviour problems. These behaviours that include aggressive behaviours and conduct problems are less susceptible to change, showing no significant decline over the intervention period. However the p-value

($p=0.077$) is of the same order as that required for the 5% level of significance ($p<0.05$), and this along with the small effect size ($ES=0.45$) above the average value of 0.4 for educational interventions, is evidence of a substantial improvement over the intervention period, which is sustained at the one-year follow-up point. Overall therefore there is compelling evidence that the good effects of groupwork are not only substantial over the intervention period, but also persist at one year according to the children.

It is however puzzling that the children's problem questionnaire (YSR) was examined in Hypothesis 1 and there was no significant improvement found over the intervention period. The explanation offered for this is that the population is different. The numbers of children present at time points T1, T2 and T3 for the waiting list data are small ($n=44$) relative to those present at the three time points T2, T3 and T4 ($n=77$). The numbers represent the organisational and methodological difficulties in making sure that all children received their first assessment 3-months prior to intervention. As these numbers show, there was a substantial portion whose first assessment was at T2 – immediately prior to intervention. These children regard their behaviour to have improved over the intervention period – significantly. The effect sizes for the period T2-T3 in Hypothesis 2A are comparable with those of Hypothesis 1A, both being small at around 0.4.

However, according to the findings for Hypothesis 1, it may be that children will show a dramatic improvement following their first assessment regardless of whether an intervention is in place or not. It might be that the findings of Hypothesis 2A that show a significant improvement T2-T3 are reflecting this effect of selection, proposed as an interpretation of Hypothesis 1A. Both hypotheses show a period of sustained improvement, or at least no significant deterioration following initial gains and there are no data available to distinguish between this so-called "selection-effect" and the effects of intervention.

Attrition:

The number of children filling in forms at all three time points is 77. This is 60 % of the sample and as it registers children still in the project after 18 months (from the initial assessment at T1) this is considered acceptable.

Hypothesis 2b

Under this hypothesis we report findings from the supplementary self-assessment measure, the Multi-Dimensional self-concept Scale (MSCS) that relate to the persistence of change over a one-year follow-up period. Hypothesis 2B examines the research question by looking at the responses of children filling in the MSCS questionnaire at time points T2, T3 and T4.

Experimental Hypothesis 2b

“According to child self-report (MSCS), the gains on self-esteem measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).”

Null Hypothesis 2B

“According to child self-report (MSCS), the gains on self-esteem measures over the intervention period (T2 – T3) are not sustained over the follow-up period (T3 – T4).”

Findings for Hypothesis 2b:

Table 2b Descriptives

MSCS	t	Mean	SD	N	do
Total T-Score	2	51.96	8.66	75	2
	3	49.12	9.44	75	
	4	46.68	8.59	75	
Academic T- Score	2	51.34	9.34	76	2
	3	49.25	10.35	76	
	4	47.34	11.49	76	
Affect T-Score	2	51.96	8.53	75	2
	3	49.04	8.82	75	
	4	46.65	9.25	75	
Competence T-Score	2	54.68	9.32	76	2
	3	51.45	9.44	76	
	4	49.36	9.49	76	
Family T-Score	2	50.42	8.74	76	2
	3	49.39	9.62	76	
	4	47.18	7.90	76	
Physical T- Score	2	50.56	8.87	75	2
	3	48.44	9.37	75	
	4	48.53	8.91	75	
Social T-Score	2	52.17	9.27	75	2
	3	49.13	9.85	75	
	4	46.52	10.71	75	

Table 2b Analysis

MSCS	Repeated Measures ANOVA & Effect Size (ES)								
	T2-T4			T2-T3			T3-T4		
	F	p<	ES	F	p<	ES	F	p<	ES
Total T-Score	15.408	0.01 (0.000)	0.61	10.376	0.01 (0.002)	0.31	6.206	0.05 (0.015)	0.27
Academic T-Score	5.849	0.01 (0.004)	0.39	4.141	0.05 (0.045)	0.21	2.145	ns (0.147)	0.18
Affect T-Score	12.972	0.01 (0.000)	0.6	11.750	0.01 (0.001)	0.34	4.819	0.05 (0.031)	0.26
Competence T-Score	13.592	0.01 (0.000)	0.57	12.800	0.01 (0.001)	0.35	4.247	0.05 (0.043)	0.22
Family T-Score	4.746	0.01 (0.01)	0.39	1.002	ns (0.315)	0.11	3.615	ns (0.061)	0.25
Physical T-Score	2.576	ns (0.080)	0.23	4.095	0.05 (0.047)	0.23	0.008	ns (0.930)	-0.01
Social T-Score	13.511	0.01 (0.000)	0.57	8.561	0.01 (0.005)	0.32	4.997	0.05 (0.028)	0.25

Reporting Results for Hypothesis 2b:

For Periods T2 – T3 & T3 – T4

Total Problem T-scores

There is a statistically significant improvement over the intervention period (T2 – T3) on this measure ($p=0.002$), mean scores falling T2-T3: 51.96 – 49.12, with a small effect size. During the follow-up period (T3 – T4) the gains continue to establish themselves, mean scores falling further 49.12 – 46.68. This improvement is statistically significant ($p=0.015$), but the effect size is small.

Academic Dimension T-Scores

There is statistically significant improvement over the intervention period on this measure ($p=0.045$), mean scores falling T2-T3: 51.34 – 49.25. This decline in mean scores continues over the follow-up period (T3 – T4), this is not statistically significant ($p=0.147$) and the effect size is small.

Affect Dimension T-scores

There is statistically significant improvement over the intervention period (T2 – T3) on this measure ($p=0.001$), mean scores falling T2-T3: 51.96 – 49.04, with a small effect size. During the follow-up period (T3 – T4) the gains continue to establish themselves mean scores falling further 49.04 – 46.65. This improvement is statistically significant ($p=0.031$), but the effect size is small.

Competence Dimension T-scores

There is statistically significant improvement over the intervention period (T2 – T3) on this measure ($p=0.001$), mean scores falling T2-T3: 54.68 – 51.45, with a small effect size. During the follow-up period (T3 – T4) the gains continue to establish themselves mean scores falling further 51.45 – 49.36. This improvement is statistically significant ($p=0.043$), but the effect size is small.

Family Dimension T-scores

There is no significant improvement over the intervention period on this measure ($p=0.315$), mean scores are falling T2-T3: 50.42 – 49.39, but the effect size is small. Although this pattern is repeated in the follow-up period, the fall in mean scores is greater 49.39 – 47.18, and statistically significant ($p=0.061$) with an increasing effect size still small at 0.25.

Physical Dimension T-scores

There is a statistically significant improvement over the intervention period on this measure ($p=0.047$), mean scores falling T2-T3: 50.56 – 48.44. This gain in mean scores is reversed by the smallest of margins over the follow-up period (T3 – T4), this is not statistically significant ($p=0.930$) and the effect size is small.

Social Dimension T-scores

There is statistically significant improvement over the intervention period (T2 – T3) on this measure ($p=0.005$), mean scores falling T2-T3: 52.17 – 49.13, with a small effect size. During the follow-up period (T3 – T4) the gains continue to establish themselves mean scores falling further 49.13 – 46.52. This improvement is statistically significant ($p=0.028$), but the effect size is small.

Summary:

The total score shows highly significant improvement over the intervention period ($p<0.01$) followed by significant improvement over the one-year follow-up period ($p<0.05$). All but the “Family” sub-scale show significant improvement T2-T3 and all show sustained improvement at one-year follow-up, with none showing significant deterioration.

On this evidence it is possible to reject the null hypothesis and to accept the

Experimental Hypothesis 2b:

“According to child self-report (MSCS), the gains on self-esteem measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).”

Interpretation of Data and Discussion of Hypothesis 2b:

Children clearly rate their own self-esteem as improving over the intervention period and continuing to improve over the following year. This is an encouraging finding for the long-term effectiveness of small groupwork. Only the “Family” sub-scale bucks this general trend, with no significant gains T2-T3 and no significant change T3-T4. There is however a continual and substantial decline in mean score over T2-T4 (T2-T3-T4 = 50.42 – 49.39 – 47.18). The lack of a dramatic drop over intervention decreases the probability of finding significance, but the overall improvement T2-T4 is found to be highly significant ($p=0.01$). Overall therefore there is compelling evidence that the good effects of groupwork are not only substantial over the intervention period, but also persistent at one-year according to the children. Indeed the MSCS clearly shows children not only holding gains made over intervention, but continuing to improve overall and according to the majority of sub-scales.

The puzzle with these findings is a similar puzzle as expressed in the interpretation for Hypothesis 2a. The MSCS was examined in Hypothesis 1B and there was no significant improvement found over the intervention period. The most obvious explanation is again that the population is different. The numbers of children present at time points T1, T2 and T3 are small ($n=27$) relative to those present at the three time points T2, T3 and T4: ($n=75$). The numbers represent the organisational and methodological difficulties in making sure that all children received their first assessment 3-months prior to intervention, as these numbers show a substantial portion whose first assessment was at T2 – immediately prior to intervention. The effect sizes are particularly relevant for sample sizes below $n=40$ and in Hypothesis 1B for the period T2-T3 they support the general view of low impact by being consistently of a trivial order. The effect sizes over the same time period in Hypothesis 2B are of a small order, certainly not showing magnitude of effect out of the ordinary, and in this latter case the change is statistically significant, i.e. unlikely to have come about by chance.

However, the “selection-effect” argument applies here as it did to Hypothesis 2a. According to the findings for Hypothesis 1b, children may show a dramatic improvement following their first assessment regardless of whether an intervention is in place or not. It might be that the findings of Hypothesis 2b that show a significant improvement T2-T3 are reflecting this effect of selection proposed as an interpretation of Hypothesis 1b. Both sets of scores show a period of sustained improvement, or at least no significant deterioration following initial gains and there is no data available to distinguish between this so-called “selection-effect” and the effects of intervention. Hypothesis 2b adds something to the interpretation offered so far. The improvements seen over the intervention period are not only sustained over the follow-up period but significant improvements continue in total scores in the follow-up period with Affect, Competence, Social and Family sub-scales just beneath levels of significance. This trend is not seen in the data relating to Hypothesis 1b, where there is no significant change in scores following the start of intervention. Indeed the mean-scores show a slight deterioration. It would appear that the trend in the data for Hypothesis 2b is a different one than that shown in Hypothesis 1b attributed in that case to a “selection effect”, it is possible therefore to propose that the gains made and sustained in hypothesis 2b and by inference 2a are not fully accounted for by “selection-effects” and this may be

considered as evidence for real and lasting treatment effects up to one-year post intervention.

Multiple Testing:

Using logarithms developed by Sakoda Cohen and Beal (1954) it is possible to predict the probability of obtaining n or more significant statistics out of N calculated statistics. In this case, the number of calculated statistics is the number of scores is six (MSCS total score not considered independent of the sub-scale scores), that are tested over three time periods giving a total of 18 calculated statistics. Of these statistics calculated for Hypothesis 2b, 16 are significant, some at $p < .05$, some at $p < .01$. Using the developed logarithms, the chance probability of obtaining this number of significant statistics is well under .001.

Attrition:

The number of children filling in forms at all three time points is 77. This is 60% of the sample and as it registers children still in the project after 18 months (from the initial assessment at T1) this is considered acceptable.

Hypothesis 2c

This section reports findings from the Achenbach teacher questionnaire, the Teacher Report Form (TRF) that relate to the persistence of change over a one-year follow-up period. Hypothesis 2C examines the research question by looking at the responses of teachers filling in the TRF questionnaire at time points T2, T3 and T4.

Experimental Hypothesis 2c

“According to teacher report (TRF), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).”

Null Hypothesis 2c

“According to teacher report (TRF), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are not sustained over the follow-up period (T3 – T4).”

Findings for Hypothesis 2c:

Table 2c Descriptives

TRF	t	Mean	SD	N	df
Total T-Score	2	56.94	10.96	104	2
	3	52.42	10.11	104	
	4	50.04	10.65	104	
Internal T- Score	2	56.01	12.46	104	2
	3	52.79	11.92	104	
	4	48.55	8.33	104	
External T-Score	2	56.78	10.93	104	2
	3	52.25	10.80	104	
	4	51.93	10.77	104	

Table 2c Analysis

TRF	Repeated Measures ANOVA & Effect Size (ES)								
	T2-T4			T2-T3			T3-T4		
	F	p<	ES	F	p<	ES	F	p<	ES
Total T-Score	22.835	0.01 (0.000)	0.64	36.162	0.01 (0.000)	0.43	4.331	0.05 (0.040)	0.23
Internal T-Score	21.041	0.01 (0.000)	0.73	11.009	0.01 (0.001)	0.26	12.654	0.01 (0.001)	0.43
External T-Score	12.775	0.01 (0.000)	0.45	24.554	0.01 (0.000)	0.42	0.066	ns (0.798)	0.03

Reporting Results for Hypothesis 2c:

For Intervention Period T2 - T3 & Follow-up period T3 – T4

Total Problem T-scores

There is statistically significant Improvement over the intervention period on this measure (p=0.000), mean scores falling T2-T3: 56.94 – 52.42. The small effect size is 0.43 of a standard deviation. There is statistically significant improvement over the follow-up period on this measure (p=0.040), mean scores falling again T3-T4: 52.42 – 50.04. The small effect size is 0.23 of a standard deviation.

Internalising Problem T-Scores

There is statistically significant improvement over the intervention period on this measure ($p=0.001$), mean scores falling T2-T3: 56.01 – 52.79. The small effect size is 0.26 of a standard deviation. There is statistically significant improvement over the follow-up period on this measure ($p=0.001$), mean scores falling T3-T4: 52.79 – 48.55. The small effect size is 0.43 of a standard deviation.

Externalising Problem T-scores

There is statistically significant improvement over the intervention period on this measure ($p=0.000$), mean scores falling T2-T3: 56.78 – 52.25. The small effect size is 0.42 of a standard deviation, above the average value for educational interventions. There is no statistically significant improvement over the follow-up period on this measure ($p=0.798$), mean scores falling T3-T4: 52.25 – 51.93. The trivial effect size is 0.03 of a standard deviation.

Summary:

The total scores show significant improvement ($p<0.01$) over the intervention period that are sustained or continued in the follow-up period. This trend is followed for Internalising and Externalising scales. On this evidence it is possible to reject the null hypothesis

Experimental Hypothesis 2c:

“According to teacher report (TRF), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).”

Interpretation of Data and Discussion of Hypothesis 2C:

The teachers of the children involved in groupwork provide compelling evidence for its effectiveness and for the persistence of gains made over the intervention period up to one year post intervention. In fact, like the children themselves reporting on their own improvements in Hypothesis 2b, they see gains continuing one-year post intervention. Externalising scores show gains made still robust at one year but not improving. Overall these findings provide support for small group interventions.

Teachers may be subject to a selection bias themselves and it is possible that focussing attention on these children might cause them to be rated differently. It is not possible to logically predict the direction of that bias with any great certainty, and in addition the teachers filling in the questionnaire at the follow-up point are different to those completing it at previous points in every case, the children having moved class-group and usually schools in the intervening period. There is therefore no way that the teachers could be aware of the previous scores for children and no reason that these scores should not be taken as indications of real and lasting change.

Attrition:

The number of children present at time points T2, T3 and T4 that the teachers have filled out TRF questionnaires on is 104. This is 80% of the research sample.

Hypothesis 2d

Under this hypothesis we report findings from the Achenbach parent questionnaire, the Child Behaviour Checklist (CBC) that relate to the persistence of change over a one-year follow-up period. Hypothesis 2d examines the research question by looking at the responses of parents filling in the CBC questionnaire at time points T2, T3 and T4.

Experimental Hypothesis 2d

“According to parent report (CBC), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).”

Null Hypothesis 2d

“According to parent report (CBC), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are not sustained over the follow-up period (T3 – T4).”

Findings for Hypothesis 2d:

Table 2d Descriptives

CBCL	t	Mean	SD	N	df
Total T-Score	2	54.18	10.17	60	2
	3	51.33	10.26	60	
	4	53.80	10.44	60	
Internal T- Score	2	53.50	10.56	60	2
	3	51.55	10.39	60	
	4	51.37	10.20	60	
External T-Score	2	52.65	9.73	60	2
	3	50.80	9.26	60	
	4	53.72	10.57	60	

Table 2d Analysis

CBCL	Repeated Measures ANOVA & Effect Size (ES)								
	T2-T4			T2-T3			T3-T4		
	F	p<	ES	F	p<	ES	F	p<	ES
Total T-Score	4.047	0.05 (0.020)	0.04	8.698	0.01 (0.005)	0.28	4.77	0.05 (0.033)	-0.24
Internal T-Score	2.066	ns (0.131)	0.21	2.703	ns (0.106)	0.19	0.023	ns (0.879)	0.01
External T-Score	3.262	0.05 (0.042)	-0.11	3.667	ns (0.060)	0.2	5.601	0.05 (0.021)	-0.3

Reporting Results for Hypothesis 2d:

For Intervention Period T2-T3 & Follow-up Period T3 – T4

Total Problem T-scores

There is statistically significant improvement over the intervention period on this measure (p=0.005), mean scores falling T2-T3: 54.18 – 51.33. The small effect size is 0.28 of a standard deviation. There is statistically significant deterioration over the follow-up period on this measure (p=0.03), mean scores rising T3-T4: 51.33 – 53.80. The small negative effect size is -0.24 of a standard deviation.

Internalising Problem T-Scores

There is no statistically significant improvement over the intervention period on this measure ($p=0.106$), mean scores falling T2-T3: 53.50 – 51.55. The small effect size is 0.19 of a standard deviation. There is no statistically significant change over the follow-up period on this measure ($p=0.879$), mean scores falling T3-T4: 51.55 – 51.37. The trivial effect size is 0.01 of a standard deviation.

Externalising Problem T-scores

There is no statistically significant improvement over the intervention period on this measure ($p=0.060$), mean scores falling T2-T3: 52.65 – 50.80. The small effect size is 0.20 of a standard deviation. There is a statistically significant deterioration over the follow-up period on this measure ($p=0.021$), mean scores rising T3-T4: 50.80 – 53.72. The small negative effect size is -0.30 of a standard deviation.

Summary:

The total problem scores as reported by the parents show highly significant ($p<0.01$) improvement over T2-T3, the intervention period, followed by significant deterioration ($p<0.05$) over the follow-up period T3-T4. Neither of the sub-scales show significant gains over the intervention period, and the Externalising scores show a significant ($p<0.05$) deterioration in follow-up. On this evidence the experimental hypothesis is rejected and the null hypothesis accepted:

“According to parent report (CBC), the gains on emotional and behavioural measures over the intervention period (T2 – T3) are not sustained over the follow-up period (T3 – T4).

Interpretation of Data and Discussion of Hypothesis 2d:

Looking at the total scores it is reasonably clear that the parents see an improvement in their children's behaviour, but over the next year the improvements do not last. The deterioration does not take the levels of disturbance back to the levels at which they started (mean scores for T2= 54.18 and for T4=53.80) but the pattern is quite clear. This trend is exaggerated in the Externalising scores in which it is possible to see the scores deteriorating over T3-T4 to the extent that mean scores at T4 are worse than at T2. The Internalising scores show a gradual decline T2-T4 that is not significant at any point, but it does provide a counterfoil to the predominant trends in the data of 2d. It shows that

there is a trend in the way the parents evaluate the children's emotional symptoms that is slightly different to how they assess the more aggressive symptoms.

These findings are not surprising. Cross-contextual gains are rare in intervention studies, and are the prime concern for clinic-based interventions. For this study, based as it is in schools, gains specific to the context are useful in themselves and the child and parent data are encouraging in this respect, providing enough outcome to justify the intervention. The parents may be reflecting changes in their child's behaviour over the period of study, or it may be a reflection of their attitude towards the child's behaviour. To say it is a mixture of the two is close to truism but it is worth wondering which factor is at play in deciding the direction of the change in scores.

The parents give evidence of an initial overall decline in problem behaviours, not significant in either of the sub-scales, followed by a return to previous levels of disturbance. This is unusual, especially when considered alongside the data from the children that show stability in gains made over the intervention period. In the examination of Hypothesis 1 parents rated children's Externalising problems as improving at the same time as children rated their own Internalising problems as improving. It is possible that this Internalising improvement is characterised initially by reflective behaviour, hence the registration by the parents of behavioural gains. However as these improvements continue, the reflective behaviour is replaced by confident and positively challenging behaviour – behaviour likely to be viewed as disruptive by the parents who do not have inside knowledge of the process that takes place in the school environment. Indeed it is a brave parent and a brave teacher that sees beyond confidence masquerading as quiet submission to an encouragement of informed inquisitiveness and positive challenge. In this respect it is noted that the Family sub-scale of the MSCS shows no significant improvement over either of the time periods, although there are signs that improvement is creeping in during T3-T4. (mean scores T2-T3-T4: 50.42 – 49.39 – 47.18, $p=0.061$: T3-T4). The children see relations at home not improving in the same way as other aspect of their lives. Over follow-up, parents report children's behaviour as relapsing from previous gains, while children rate familial relations as gradually improving

Attrition:

The number of parent filling questionnaires on children at all three time points is 60. This is 40% of the research sample and is low, but in keeping with parent response from similar studies (see e.g. (Kolvin, Garside, Nicol, Macmillan, Wolstenholme, & Leitch 1981). There is no evidence that these parents have self-selected, but it is likely that those with least empathy with the research aims will have resisted the attempts of the research team to contact them. It is therefore possible to speculate that if the sample had been greater in number the findings might have been less positive.

Section 5.3: Hypothesis 3

Question 3. Do gains in emotional/behavioural profile depend on what goes on in the group?

If you compare a group who are engaged in child-focussed developmental and therapeutic activities (AGI) with a group engaged in curriculum tasks (CGI), will you notice a difference in emotional/behavioural profile?

Is there any difference between the groups in the short-term? To answer this the impact on the AGI group can be compared to the impact on the CSG group the intervention period T2-T3. The groups are experimentally controlled for variables such as time of day, duration, facilitator and general child-centred approach, but are different in the content of the sessions delivered. Because the AGI sessions are more clearly focussed on the enrichment of relationships and the resolution of conflict situations, it is predicted that the AGI group will be more effective in reducing problem behaviours and increasing self-esteem.

Hypothesis 3:

According to parents, teachers and self-report, the groupskills intervention (AGI) will be more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems and promoting self-esteem.

The full form of this hypothesis is given below:

- a) According to child self-report (YSR), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).**
- b) According to child self-report (MSCS), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting self-esteem over the intervention period (T2 –T3).**
- c) According to teacher report (TRF), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).**

- d) According to parent report (CBC), the groupskills intervention (AGI) is more effective than the small-group curriculum studies Intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).

This hypothesis set is concerned with the intervention period T2-T3. Instead of looking at the whole intervention group (AGI + CSG) these hypotheses look for a difference in gains between the groups. Gains are measures by Achenbach scales and the Multi-dimensional Self-Concept Scale as in Hypotheses 1 and 2, and as in these prior hypotheses, the analyses applied are the repeated measures ANOVA and effect size to measure the magnitude of change.

Hypothesis 3a

Experimental Hypothesis 3a

“According to child self-report (YSR), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).”

Null Hypothesis 3a

“According to child self-report (YSR), there is nothing to distinguish between the groupskills intervention (AGI) and the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).”

Findings for Hypothesis 3a

Table 3a Descriptives

YSR	t	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Total T-Score	2	55.38	12.16	45	53.07	12.26	43	1
	3	52.09	11.72	45	49.20	10.29	43	
Internal T-Score	2	55.04	11.70	45	53.74	11.75	43	1
	3	50.93	12.07	45	49.19	9.91	43	
External T-Score	2	53.91	12.21	45	52.58	12.16	43	1
	3	52.31	11.31	45	50.30	10.24	43	

Table 3a Analysis

YSR	Repeated measures ANOVA		Effect Size (ES)	
	Time by Condition (AGI vs CSG)		T2-T3	
	F	p<	AGI	CSG
Total T-Score T2-T3	0.49	ns (0.825)	0.28	0.34
Internal T- Score T2-T3	0.031	ns (0.861)	0.35	0.42
External T-Score T2-T3	0.081	ns (0.776)	0.14	0.2

Reporting Results for Hypothesis 3a:
Total Problem Scores

For groups compared to each other (AGI vs CSG)
There is no statistically significant difference reported by YSR between the groups over intervention period. AGI scores falling 55.38 – 52.09 with a small effect size of 0.28, and CSG scores falling 53.07 – 49.20 with a small effect size of 0.34.

Internalising Problem Scores

For groups compared to each other (AGI vs CSG)
There is no statistically significant difference reported by YSR between the groups over intervention period. AGI scores falling 55.04 – 50.93 with a small effect size of 0.35, and CSG scores falling 53.74 – 49.19 with a small effect size of 0.42.

Externalising Problem Scores

For groups compared to each other (AGI vs CSG)
There is no statistically significant difference reported by YSR between the groups over intervention period. AGI scores falling 53.91 – 52.31 with a trivial effect size of 0.14, and CSG scores falling 52.58 – 50.30 with a trivial effect size of 0.2

Summary:

The trend is for child reported improvement over the intervention period, a finding supported by Hypothesis 2a, however, there is no evidence on this measure that distinguishes between the groups. As far as the children are concerned it makes no difference to their problems whether they do the AGI or CSG. On this evidence the null hypothesis is accepted:

“According to child self-report (YSR), there is nothing to distinguish between the groupskills intervention (AGI) and the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).”

Interpretation of Data and Discussion of Hypothesis 3a:

The message from the children reporting on the YSR is unequivocal: they do not distinguish between the groups they attend in the short-term. The sample size is cut to 45 and this will reduce the chances of findings statistically significant differences between the groups, but effect sizes report magnitude of change regardless of sample size and when these scores for the AGI and CSG groups are compared there is no substantive difference between them. The Internalising score favours the CSG group the Externalising favouring the AGI group, but these differences are not large enough to register as a difference in magnitude.

Attrition:

The number of children responding is 45 for the AGI group and 43 for the CSG group, giving a total of 88. This is 70% of the research sample and there does not appear to be any effects of attrition that are affecting one intervention condition over another.

Hypothesis 3b

Experimental Hypothesis 3b

“According to child self-report (MSCS), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting self-esteem over the intervention period (T2 –T3).”

Null Hypothesis 3B

“According to child self-report (MSCS), there is nothing to distinguish between the groupskills intervention (AGI) and the small-group curriculum studies intervention (CSG) in promoting self-esteem over the intervention period (T2 –T3).”

Findings for Hypothesis 3b:

Table 3b Descriptives

MSCS	T	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Total T-Score	2	51.62	9.17	42	52.61	8.60	41	1
	3	49.17	10.65	42	49.80	8.02	41	
Academic T-Score	2	51.24	9.49	42	52.46	9.56	41	1
	3	49.48	12.12	42	49.76	8.01	41	
Affect T-Score	2	51.81	9.38	42	52.02	7.92	41	1
	3	48.24	9.88	42	50.00	7.69	41	
Compet-ence T-Score	2	54.71	8.93	42	54.43	9.67	41	1
	3	50.57	10.63	42	52.44	7.89	41	
Family T-Score	2	48.31	9.62	42	53.20	7.28	41	1
	3	49.24	9.86	42	49.93	9.30	41	
Physical T-Score	2	49.62	8.40	42	51.71	9.52	41	1
	3	48.33	10.65	42	48.66	7.78	41	
Social T-Score	2	53.76	9.37	42	50.18	9.19	41	1
	3	50.17	10.46	42	48.70	8.89	41	

Table 3b Analysis T2-T3

	Repeated measures ANOVA		Effect Size (ES)	
	Time by Condition (AGI vs CSG)		T2-T3	
	F	p<	AGI	CSG
Total T-Score	0.43	ns (0.837)	0.25	0.34
Academic T-Score	0.225	ns (0.636)	0.16	0.31
Affect T-Score	0.881	ns (0.351)	0.37	0.26
Compet -ence T-Score	1.515	ns (0.222)	0.41	0.23
Family T-Score	4.987	0.05 (0.028)	-0.1	0.4
Physica l T-Score	0.800	ns (0.374)	0.14	0.35
Social T-Score	1.137	ns (0.289)	0.36	0.16

Reporting Results for Hypothesis 3b:

Total Problem T-scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported by MSCS between the groups over intervention period, AGI scores falling 51.62 – 49.17 with a small effect size of 0.25, and CSG scores falling 52.61 – 49.80 with a small effect size of 0.34.

Academic Dimension T-Scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported by MSCS between the groups over intervention period, AGI scores falling 51.24 – 49.48 with a small effect size of 0.16, and CSG scores falling 52.46 – 49.76 with a small effect size of 0.31.

Affect Dimension T-scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported by MSCS between the groups over intervention period, AGI scores falling 51.81 – 48.24 with a small effect size of 0.37, and CSG scores falling 52.02 – 50.00 with a small effect size of 0.26.

Competence Dimension T-scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported by MSCS between the groups over intervention period, AGI scores falling 54.71 – 50.57 with a small effect size of 0.41, and CSG scores falling 54.43 – 52.44 with a small effect size of 0.23.

Family Dimension T-scores

For groups compared to each other (AGI vs CSG)

There is statistically significant difference reported by MSCS between the groups over intervention period ($p=0.028$), AGI scores rising 48.31 – 49.24 with a trivial effect size of 0.10, and CSG scores falling 53.20 – 49.93 with a small effect size of 0.40

Physical Dimension T-scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported by MSCS between the groups over intervention period, AGI scores falling 49.62 – 48.33 with a small effect size of 0.14, and CSG scores falling 51.71 – 48.66 with a small effect size of 0.35.

Social Dimension T-scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported by MSCS between the groups over intervention period, AGI scores falling 53.76 – 50.17 with a small effect size of 0.36, and CSG scores falling 50.18– 48.70 with a small effect size of 0.16.

Summary:

The children report significant gains in self-esteem as a result of being involved in small groupwork, but give little statistically significant evidence for distinguishing between the AGI and the CSG. The family sub-scale reports a significant ($p < .05$) difference between the groups. The mean scores show the AGI group reporting the family situation deteriorating, while the CSG group reports it as improving. The effect sizes are more illuminating and unaffected by sample size reveal some interesting effects of the different interventions. These effects favour the AGI on 3 sub-scales and the CSG on 3 sub-scales and therefore give no evidence to distinguish overall between the groups. For these reasons the null hypothesis is accepted:

“According to child self-report (MSCS), there is nothing to distinguish between the groupskills intervention (AGI) and the small-group curriculum studies intervention (CSG) in promoting self-esteem over the intervention period (T2 –T3).”

Interpretation of Data and Discussion of Hypothesis 3b:

In reporting the impact of small groupwork on their self-image the children are similarly unequivocal as they were in reporting their problems. There is little statistically significant evidence to differentiate between the groups. The anomalous finding here is the family sub-scale, which reports change in the opposite direction to that predicted: it favours the CSG group. Speculation as to the reasons for this focus on the CSG children reporting an easier time at home because they have something to say to their parents when they ask about progress at school. The CSG children are able to talk about their involvement in the curriculum group, something parents will understand easily, while AGI children will talk about creative expressive activities that are likely to provoke more questioning.

The effect size scores do show discernable differences, with the CSG group reporting differential gains of family scores, as discussed above, academic scores which might be expected and physical scores, which is more challenging to explain. The AGI children score more highly on affect, competence and social sub-scales, which is to be expected

from the nature of the intervention. However these differences are within the same category (small) of effect size, and they are evenly distributed in terms of the intervention condition favoured.

Multiple Testing:

Using logarithms developed by Sakoda Cohen and Beal (1954) it is possible to predict the probability of obtaining n or more significant statistics out of N calculated statistics. In this case, the number of calculated statistics is the number of scores is seven (MSCS total score plus six sub-scale scores), that are tested over one time period (T2 – T3) for two groups (AGI +CSG and AGI vs CSG) giving a total of 14 calculated statistics. Of these statistics calculated for Hypothesis 3b, 7 are significant, some at $p < .05$ some at $p < .01$. Using the developed logarithms, the chance probability of obtaining this number of significant statistics is well under .001.

Attrition:

The number of children responding is 42 for the AGI group and 41 for the CSG group, giving a total of 83. This is 65% of the research sample and there does not appear to be any effects of attrition that are affecting one intervention condition over another.

Hypothesis 3c

Experimental Hypothesis 3c

“According to teacher report (TRF), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).”

Null Hypothesis 3c

“According to teacher report (TRF), there is nothing to distinguish between the groupskills intervention (AGI) and the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).”

Findings for Hypothesis 3c:

Table 3c Descriptives

TRF	T	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Total T-Score	2	58.57	10.48	58	55.15	11.25	59	1
	3	51.52	8.89	58	53.49	10.65	59	
Internal T-Score	2	57.26	12.59	58	54.49	12.01	59	1
	3	51.17	10.55	58	53.93	12.54	59	
External T-Score	2	58.41	11.90	58	55.59	10.39	59	1
	3	52.31	8.52	58	52.93	12.34	59	

Table 3c Analysis T2-T3

TRF	Repeated measures ANOVA		Effect Size (ES)	
	Time by Condition (AGI vs CSG)		T2-T3	
	F	p<	AGI	CSG
Total T-Score	15.210	0.01 (0.000)	0.73	0.15
Internal T-Score	10.086	0.01 (0.000)	0.53	0.05
External T-Score	4.041	0.05 (0.047)	0.61	0.25

Reporting Results for Hypothesis 3c:

Total Problem Scores

For groups compared to each other (AGI vs CSG)

There is a highly significant difference reported by TRF between the groups over intervention period ($p=0.00$), AGI scores falling 58.57 – 51.52 with a moderate effect size of 0.73, and CSG scores falling 55.15 – 53.49 with a small effect size of 0.15.

Internalising Problem Scores

For groups compared to each other (AGI vs CSG)

There is a highly significant difference reported by TRF between the groups over intervention period (0.000), AGI scores falling 57.26 – 51.17 with a small effect size of 0.53, and CSG scores falling 54.49 – 53.93 with a trivial effect size of 0.05.

Externalising Problem Scores

For groups compared to each other (AGI vs CSG)

There is a significant difference reported by TRF between the groups over intervention period ($p=0.047$), AGI scores falling 58.41 – 52.31 with a moderate effect size of 0.61, and CSG scores falling 55.59 – 52.93 with a small effect size of 0.25.

Summary:

The teachers report a significant difference between the groups, the AGI out-performing the CSG in gains made on total, Internalising and Externalising scores. The magnitude of change as reported by the effect size analysis reveals small and moderate change in the AGI group, well above the average value for educational interventions, compared to small and trivial change in the CSG group. On this evidence the null hypothesis is rejected,

Experimental Hypothesis 3c.

“According to teacher report (TRF), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).”

Interpretation of Data and Discussion of Hypothesis 3C:

The teachers are as unequivocal in the distinction they draw between the groups as the children are in their lack of distinction. Teachers see children improving over the intervention period, reporting highly significant gains ($p=.000$), but also see the AGI improving in their emotional symptoms ($p=.000$) and also in behavioural terms ($p=.047$). This is change in the direction predicted, the nature of the AGI content being thought to promote change at a level not achieved by the more routine content of the CSG. Caution needs to be exercised before conclusions can be drawn. Firstly this finding, although encouraging, is in direct opposition to the findings from the children themselves, also the teachers at this stage are not blind to treatment status and will have had no difficulty in

discerning which intervention carried the more importance, in other words, this finding is susceptible to bias. Countering this bias is the professional standards of the teachers involved and their lack of investment in any particular outcome – why should they want the AGI group to improve more than the CSG group? Also the questionnaires are filled in random order, and the inclination to identify a child as a member of a particular group that filled a fraction of the class time over one term could be thought of as present but not significant.

Attrition:

The number of children for whom the teachers completed a questionnaire is 58 for the AGI group and 59 for the CSG group, giving a total of 117. This is 90% of the research sample and there do not appear to be any effects of attrition that are affecting one intervention condition over another.

Hypothesis 3d

Experimental Hypothesis 3d

“According to parent report (CBC), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).”

Null Hypothesis 3d

“According to parent report (CBC), there is nothing to distinguish between the groupskills intervention (AGI) and the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 –T3).”

Findings for Hypothesis 3d:

Table 3d Descriptives

CBCL	T	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Total T-Score	2	54.47	9.42	38	54.77	10.25	39	1
	3	50.39	9.80	38	52.15	10.05	39	
Internal T-Score	2	53.71	9.52	38	53.36	10.61	39	1
	3	49.18	9.28	38	52.69	10.32	39	
External T-Score	2	53.74	10.33	38	53.41	9.78	39	1
	3	51.23	9.72	38	51.30	8.67	39	

Table 3d Analysis

CBCL	Repeated measures ANOVA		Effect Size (ES)	
	Time by Condition (AGI vs CSG)		T2-T3	
	F	p<	AGI	CSG
Total T-Score	0.733	ns 0.395	0.43	0.26
Internal T-Score	3.484	ns (0.066)	0.48	0.06
External T-Score	0.053	ns (0.818)	0.25	0.23

Reporting Results for Hypothesis 3d:

Total Problem Scores

For groups compared to each other (AGI vs CSG)

There is no significant difference reported by CBC between the groups over intervention period, AGI scores falling 53.71 – 49.18 with a small effect size of 0.43, and CSG scores falling 53.36 – 52.69 with a small effect size of 0.26.

Internalising Problem Scores

For groups compared to each other (AGI vs CSG)

There is no significant difference reported by CBC between the groups over intervention period ($p=0.066$), AGI scores falling 53.71 – 49.18 with a small effect size of 0.48, and CSG scores falling 53.36 – 52.69 with a trivial effect size of 0.06.

Externalising Problem Scores

For groups compared to each other (AGI vs CSG)

There is no significant difference reported by CBC between the groups over intervention period, AGI scores falling 53.74 – 51.23 with a small effect size of 0.25, and CSG scores falling 53.41 – 51.30 with a small effect size of 0.23.

Summary:

The parents report some differential improvements in the group. The CBC report, like the TRF and YSR is in essence two separate scales that are amalgamated into a total score. The Externalising behaviours scale does not report statistically significant change, but the Internalising score is marginally significant ($p=0.06$). The effect size scores support this differential effect by reporting an AGI score of 0.48 and a CSG score of 0.06. The parents are therefore considered to distinguish between the intervention groups, favouring the AGI over the CSG on the internalising problem scale. The effect size scores for the Externalising behaviours are not really distinguishable, and the evidence is split. Because the sample size is small and because parental data are traditionally unresponsive, this finding of differential benefit is emphasised and for these reasons the null hypothesis is rejected.

Experimental Hypothesis 3d:

"According to parent report (CBC), the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in reducing behavioural and emotional problems over the intervention period (T2 – T3)."

Interpretation of Data and Discussion of Hypothesis 3d:

The encouraging findings from these data are that children improve outside the intervention context: the impact of intervention on children is powerful enough to have positive effects at home. Parents do report differential effects favouring the Internalising behaviours of the AGI children. This is in keeping with the nature of the intervention that

is building an emotional vocabulary to deal with the external situations. It will be interesting to see whether the Externalising behaviours are affected in any way in subsequent analyses.

Two trends worth noting are the Internalising score approaching significance on the time by condition ANOVA and the effect sizes consistently higher for the AGI group. The Internalising score distinguishes between the groups at a level $p=0.066$, and that is despite the parents reporting less significant gains on Internalising scores for AGI+CSG. The magnitude of change is also enhanced for the AGI group (effect size of 0.48 compared to 0.06 for the CSG – see table 3d Analysis). This finding may point to real change on Internalising measures for the AGI group: in the context of the other CBC reported scores it certainly seems anomalous. The other trend in the data is the magnitude of change in the AGI group being at a consistently higher level than the CSG, pointing towards an effect (although below the level for significant consideration) in the AGI group not present in the CSG.

Attrition:

The number of children for whom the teachers completed a questionnaire is 38 for the AGI group and 39 for the CSG group, giving a total of 77. This is 70% of the research sample and there does not appear to be any effects of attrition that are affecting one intervention condition over another.

Section 5.4: Hypothesis 4

Question 4. How do any differences in emotional/behavioural profile between the groups alter as time goes by?

If you survey the respondents at intervals after the interventions are over will you still be able to detect differences in the gains made by the different groups?

Is it possible to distinguish between the Intervention groups at a follow-up point of one year? To answer this the AGI and CSG groups are compared at one year following intervention to see if there is any difference between the scores on the questionnaires. The AGI intervention is considered to be more able to affect the multiple contexts in which an individual will express disaffection also the multiple areas of emotional and social function that need to be addressed in order to provide lasting change. It is therefore predicted that the positive impact on AGI intervention group will be greater at the follow-up point of one year (T4) than the CSG group whose intervention has focussed on using the group to achieve curriculum tasks.

This hypothesis looks for evidence of long-term benefit measured from T2 – T4. It is during this period that analysis examines how each group differently sustains benefits. The time period T3-T4 is also presented in order to reflect in more detail on the nature of any differential gains.

Hypothesis 4:

According to parents, teachers and self-report, the AGI group will show greater long-term benefits on behavioural, emotional and self-concept measures than the CSG group, measured prior to intervention and at one year.

According to child self-report (YSR), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.

According to child self-report (MSCS), the AGI group will show greater long-term benefits on self-concept measures than the CSG group, measured prior to intervention and at one year.

According to teacher report (TRF), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.

According to parent report (CBC), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.

Hypothesis 4a

Experimental Hypothesis 4a

“According to child self-report (YSR), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.”

Null Hypothesis 4A

“According to child self-report (YSR), the AGI group will show no greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.”

Findings for Hypothesis 4a:

Table 4a Descriptives

YSR	t	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Total T-Score	2	55.69	12.47	39	52.82	12.00	38	2
	3	52.08	11.87	39	48.11	9.66	38	
	4	50.36	13.23	39	47.39	11.68	38	
Internal T-Score	2	55.49	12.04	39	53.89	11.67	38	2
	3	50.92	12.36	39	48.16	9.21	38	
	4	49.31	12.76	39	47.11	10.74	38	
External T-Score	2	54.15	12.40	39	51.89	11.23	38	2
	3	52.10	11.65	39	49.26	9.15	38	
	4	51.64	12.62	39	47.76	13.36	38	

Table 4a Analysis (ANOVA)

YSR	Repeated Measures ANOVA – Time & Condition (AGI vs CSG)			
	T2-T4		T3-T4	
	F	p<	F	p<
Total T-Score	0.079	ns (0.924)	0.104	ns (0.748)
Internal T-Score	0.077	ns (0.925)	0.034	ns (0.854)
External T-Score	0.170	ns (0.844)	0.140	ns (0.709)

Table 4a Analysis (Effect Size)

YSR	Effect Size			
	T2-T4		T3-T4	
	AGI	CSG	AGI	CSG
Total T-Score	0.42	0.46	0.14	0.07
Internal T-Score	0.5	0.61	0.13	0.11
External T-Score	0.2	0.34	0.04	0.14

Reporting Results for Hypothesis 4a:
Total Problem Scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference between these groups according to the reports given by children at any time point. The mean scores of each group fall steadily, but broadly mirror each other in the rate at which they fall, the only difference being in the AGI group scoring higher in both initial scores (AGI=55.69, CSG=52.82) and the final scores (AGI=50.36, CSG=47.39). The effect sizes are of the same order for each group.

Internalising Problem Scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference between the groups reported at any point. The effect size scores are of the same order for each group.

Externalising Problem Scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference between the groups at any time point. The mean scores for AGI group start higher (AGI=54.15, CSG=51.89) but both fall by 3 over T2-T4. Effect size scores are of the same order for each group.

Summary:

There is no evidence to distinguish between AGI and CSG over the T2-T4 period. The experimental hypothesis is rejected and the null hypothesis accepted:

“According to child self-report (YSR), the AGI group will show no greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.”

Interpretation of Data and Discussion of Hypothesis 4a:

The children, when rating their problems on the YSR, do not distinguish between the AGI condition and the CSG condition. However, both conditions produce a magnitude of change, above the average value of 0.4 for educational interventions in Total and Internalising scores that is sustained at one year. According to children, it does not matter whether they have had the devised groupwork programme or the curriculum studies group, they feel their emotional problems have improved and stay improved. The behavioural difficulties they experience show a continued improvement as well with a smaller effect size (AGI= 0.2, CSG= 0.34) over the period from beginning of intervention.

Attrition:

The numbers of children completing questionnaires at all time points are 39 for the AGI group and 38 for the CSG group. This gives a total of 77 that is 60% of the research sample. There does not appear to be an effect of attrition that means one group is unfairly represented.

Hypothesis 4b

Experimental Hypothesis 4b

“According to child self-report (MSCS), the AGI group will show greater long-term benefits on self-concept measures than the CSG group, measured prior to intervention and at one year.”

Null Hypothesis 4B

“According to child self-report (MSCS), the AGI group will show no greater long-term benefits on self-concept measures than the CSG group, measured prior to intervention and at one year.”

Findings for Hypothesis 4b:

Table 4b Descriptives

MSCS	t	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Total T-Score	2	51.59	8.86	37	52.32	8.57	38	2
	3	48.78	10.86	37	49.45	7.96	38	
	4	47.16	9.65	37	46.21	7.53	38	
Academic T-Score	2	51.11	9.46	37	51.58	9.34	38	2
	3	49.42	12.43	37	49.08	7.91	38	
	4	46.00	11.61	37	48.68	11.36	38	
Affect T-Score	2	51.78	9.33	37	52.13	7.80	38	2
	3	48.19	10.02	37	49.87	7.53	38	
	4	46.70	10.52	37	46.60	7.96	38	
Compet- -ence T-Score	2	54.94	8.68	37	54.42	10.03	38	2
	3	50.61	10.62	37	52.29	8.15	38	
	4	49.74	9.80	37	48.97	9.30	38	
Family T-Score	2	47.92	9.42	37	52.92	7.30	38	2
	3	48.97	10.06	37	49.82	9.27	38	
	4	48.24	8.02	37	46.13	7.75	38	
Physica l T-Score	2	49.63	8.28	37	51.51	9.45	38	2
	3	48.11	10.76	37	48.78	7.84	38	
	4	49.13	9.52	37	47.92	8.33	38	
Social T-Score	2	54.21	9.23	37	50.08	8.95	38	2
	3	49.61	10.75	37	48.65	8.95	38	
	4	47.32	11.06	37	45.70	10.43	38	

Table 4b Analysis (ANOVA)

MSCS	Repeated Measures ANOVA – Time & Condition (AGI vs CSG)			
	T2-T4 df=1		T3-T4 df=1	
	F	p<	F	p<
Total T-Score	0.494	ns (0.611)	0.677	ns (0.413)
Academic T-Score	0.894	ns (0.411)	1.355	ns (0.248)
Affect T-Score	0.389	ns (0.678)	0.665	ns (0.418)
Competence T-Score	0.857	ns (0.426)	1.462	ns (0.230)
Family T-Score	5.886	0.01 (0.003)	0.207	0.05 (0.021)
Physical T-Score	1.093	ns (0.338)	0.805	ns (0.373)
Social T-Score	1.186	ns (0.308)	0.078	ns (0.781)

Table 4b (Effect Size)

MSCS	Effect Size			
	T2-T4		T3-T4	
	AGI	CSG	AGI	CSG
Total T-Score	0.48	0.76	0.16	0.42
Academic T-Score	0.49	0.28	0.3	0.04
Affect T-Score	0.51	0.7	0.15	0.42
Competence T-Score	0.57	0.56	0.09	0.38
Family T-Score	-0.04	0.9	0.08	0.44
Physical T-Score	0.06	0.41	-0.1	0.11
Social T-Score	0.68	0.45	0.21	0.31

Reporting Results for Hypothesis 4b:

Total MSCS T-scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported between the groups over time, and nothing approaching significance. Effect sizes favour CSG T2-T4 (CSG is moderate=0.76, AGI is small=0.48).

Academic Dimension T-Scores

For groups compared to each other (AGI vs CSG)

Scores report no statistically significant difference between the groups over time, with mean scores tracking each other in continuing improvement. Effect sizes favour the AGI group T2-T4 (AGI=0.49, CSG=0.28).

Affect Dimension T-scores

For groups compared to each other (AGI vs CSG)

No statistically significant difference between the groups reported over time with mean scores tracking each other almost exactly. Effect sizes favour CSG T2-T4 (CSG is moderate =0.7, AGI is small=0.51).

Competence Dimension T-scores

For groups compared to each other (AGI vs CSG)

No statistically significant difference between the groups reported over time with mean scores tracking each other almost exactly. Effect sizes are very similar.

Family Dimension T-scores

For groups compared to each other (AGI vs CSG)

There is a statistically significant difference reported between the groups over time ($p<0.01$), AGI mean scores T2-T3-T4= 47.92 – 48.97 – 48.24. CSG mean scores T2-T3-T4= 54.42 – 52.29 – 48.97). Effect sizes favour CSG T2-T4 (CSG is moderate =0.9, AGI is trivial =0.04 and in the negative direction).

Physical Dimension T-scores

For groups compared to each other (AGI vs CSG)

No statistically significant difference between the groups reported over time with mean scores tracking each other almost exactly. Effect sizes favour CSG T2-T4 (CSG is small $=0.41$, AGI is trivial $=0.06$).

Social Dimension T-scores

For groups compared to each other (AGI vs CSG)

No statistically significant difference between the groups reported over either time period with mean scores tracking each other almost exactly. Effect sizes favour the AGI T2-T4 (CSG is small $=0.45$, AGI is moderate $=0.68$).

Summary:

There is no statistically significant evidence to differentiate between the AGI and CSG groups at follow-up point T4 on total MSCS scores nor on any of the sub scales, with the exception of the family scale delivering $p < 0.05$ in favour of the CSG. The effect size analysis favours the CSG group on three sub-scales and the AGI on two. This evidence makes it clear that the AGI group does not have any clear benefit over the CSG group. The experimental hypothesis is rejected and the null hypothesis accepted.

“According to child self-report (MSCS), the AGI group will show no greater long-term benefits on self-concept measures than the CSG group, measured prior to intervention and at one year.”

Interpretation of Data and Discussion of Hypothesis 4b:

The ANOVA analysis shows statistically significant differences between the groups on one sub-scale, that of the family. The small sample size is a factor in the lack of significance and the effect size analysis is helpful in discerning difference between the

groups. These analyses show a trend favouring the CSG group except on the social subscale. The children, when scoring their own self-concept reveal that they find it more helpful to be in the curriculum group. The strong finding from the family sub-scale appeared in the analysis of Hypothesis 3 and the effects are detectable one-year post intervention. This finding, along with the general finding that increased self-concept is linked to curriculum studies was anticipated through consideration of the literature, but it was not predicted to be of a higher order than the gains made by the AGI group. In order to explain this, the general argument is advanced that self-concept is not always a reliable indicator of social and emotional health. Many criminals have a very high self-concept, and a strong sense of self-righteousness (Vermeiren et al 2004). One recalls the famous comment from the father of the youth killed in the car chase that sparked the Meadow Well riots in 1991. On national television he tearfully said "My son was no joy-rider. He was a car-thief." While it is right to have compassion for a grieving father desperate to salvage something from an awful situation, it is useful in the context of these findings from the MSCS to reflect that when at-risk young people report strong self concept relating to family, emotion and physical appearance they might be saying that their stolen clothes make them look good, their family is in their pocket and they do not feel sad about the hurt they cause. While it is recognised the this point cannot be pushed too far, and to undercut the concepts used in these pages to make distinctions and draw conclusions is counter-productive, it is worth recognising that there is a counter-argument, and indeed the complexity of self-esteem is recognised in recent and thorough reviews of the area: (Emler 2001).

Multiple Testing:

Using logarithms developed by Sakoda Cohen and Beal (1954) it is possible to predict the probability of obtaining n or more significant statistics out of N calculated statistics.

In this case, the number of calculated statistics is the number of scores is seven (MSCS total score plus six sub-scale scores), that are tested over three time periods for two groups (AGI + CSG and AGI vs CSG) giving a total of 42 calculated statistics. Of these statistics calculated for Hypothesis 4b, 18 are significant, some at $p < .05$ some at $p < .01$. Using the developed logarithms, the chance probability of obtaining this number of significant statistics is well under .001.

Attrition:

The numbers of children completing questionnaires at all time points are 37 for the AGI group and 38 for the CSG group. This gives a total of 76 that is 60% of the research sample. There does not appear to be an effect of attrition that means one group is unfairly represented.

Hypothesis 4c

Experimental Hypothesis 4c

“According to teacher report (TRF), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.”

Null Hypothesis 4c

“According to teacher report (TRF), the AGI group will show no greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.”

Findings for Hypothesis 4c:

Table 4c Descriptives

TRF	t	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Total T-Score	2	58.47	10.61	51	55.47	11.20	53	2
	3	51.61	9.18	51	53.21	10.96	53	
	4	50.27	8.99	51	49.81	12.12	53	
Internal T-Score	2	57.35	12.69	51	54.72	12.22	53	2
	3	51.47	10.57	51	54.06	13.05	53	
	4	48.31	7.69	51	48.77	8.98	53	
External T-Score	2	58.00	11.27	51	55.60	10.56	53	2
	3	52.25	8.50	51	52.25	12.71	53	
	4	52.02	9.89	51	51.85	11.65	53	

Table 4c Analysis (ANOVA)

	Repeated Measures ANOVA – Time & Condition (AGI vs CSG)			
	T2-T4		T3-T4	
	F	p<	F	p<
Total T-Score	2.498	ns (0.085)	0.808	ns (0.371)
Internal T-Score	2.630	ns (0.075)	0.793	ns (0.375)
External T-Score	0.771	ns (0.464)	0.004	ns (0.948)

Table 4c Analysis (Effect Size)

TRF	Effect Size			
	T2-T4		T3-T4	
	AGI	CSG	AGI	CSG
Total T-Score	0.84	0.49	0.15	0.3
Internal T-Score	0.92	0.57	0.35	0.49
External T-Score	0.57	0.34	0.03	0.03

Reporting Results for Hypothesis 4c:

Total Problem Scores

For both groups compared to each other (AGI vs CSG)

Mean scores show the AGI starting with a higher score at T2 58.47 falling to 50.27 at T4 and CSG starting at 55.47 falling to 49.81 over the same time period. There is no statistically significant difference reported between the groups over time at the 1% and 5% levels but statistical significance does show at the 10% level.

Effect sizes favour the AGI T2-T4 (CSG is small =0.49, AGI is moderate =0.84).

Internalising Problem Scores

For both groups compared to each other (AGI vs CSG)

Mean scores show the AGI starting with a higher score 57.85 falling to 48.31 and CSG starting at 54.72 falling to 48.77. There is no statistically significant difference reported between the groups over time at the 1% and 5% levels but statistical significance does show at the 10% level.

Effect sizes favour the AGI T2-T4 (effect size for CSG is small =0.57, AGI is moderate =0.92).

Externalising Problem Scores

For groups compared to each other (AGI vs CSG)

Mean scores show the AGI starting with a higher score 58.00 falling to 52.02 and CSG starting at 55.60 falling to 51.85. There is no statistically significant difference reported between the groups over the follow-up period ($p=0.948$) with mean scores for both groups staying constant.

Effect sizes favour the AGI T2-T4 (CSG is small $=0.34$, AGI is small $=0.57$) in the magnitude of change reported in this dimension.

Summary:

There is no statistically significant difference between the groups, but the sample size is small. Effect size analysis T2-T4 gives consistent indications of an effect differentially favouring the AGI group. Weighting is given to the effect size analysis because of the reduced sample size, and together with the account taken of the $p<0.1$ finding in the Internalising scale, the experimental hypothesis is considered to have received some support from these data.

Experimental Hypothesis 4c

“According to teacher report (TRF), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.”

Interpretation of Data and Discussion of Hypothesis 4c:

The teachers report no statistically significant differential benefit for each group at the 1% and 5% levels. Although the sample size is respectable at 80% of the research population, the difference in gains is marginal and the numbers insufficient to give the desired p -value. The difference in gains is reported as marginal for a number of reasons: firstly the group intervention effect is strong and can be detected at T4 as was discussed (and celebrated) in Hypothesis 2. Secondly, the teachers are all blind to treatment status with no knowledge of the research aims. This is a strength of the data at this point, but it does mean that the teachers are subject to variability in areas such as which period to consider children over when reporting change. Statistical significance does show though at the 10% level ($p=0.075$) for the Internalising scale and this is reflected in the Total scale ($p=0.085$). For reasons discussed above this is considered a positive finding of some import. The strength of the effect size is the independence of analysis from sample size. In this case there is a good argument for giving additional weight to effect size,

particularly as the direction of reported difference is unequivocal. The result is a detectable difference between the groups in the school environment one-year post intervention using effect size analysis, supported by statistical significance ($p < 0.1$) on the Internalising scale.

Attrition:

The numbers of teacher-completed questionnaires for children at all time points are 51 for the AGI group and 53 for the CSG group. This gives a total of 104 that is 80% of the research sample. There does not appear to be an effect of attrition that means one group is unfairly represented.

Hypothesis 4d

Experimental Hypothesis 4d

“According to parent report (CBC), the AGI group will show greater long-term benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.”

Null Hypothesis 4d

“According to parent report (CBC), the AGI group will show no greater benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year.”

Findings for Hypothesis 4d:

Table 4d Descriptives

	t	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Total T-Score	2	53.68	10.07	28	54.63.	10.39	32	2
	3	50.11	10.62	28	52.41	9.98	32	
	4	54.39	11.00	28	53.28	10.07	32	
Internal T-Score	2	53.75	10.60	28	53.28	10.70	32	2
	3	49.07	10.48	28	53.72	9.97	32	
	4	52.29	11.56	28	50.56	8.95	32	
External T-Score	2	51.89	9.84	28	53.31	9.733	32	2
	3	50.46	10.01	28	51.09	8.70	32	
	4	53.36	10.34	28	54.03	10.91	32	

Table 4d Analysis (ANOVA)

	Repeated Measures ANOVA – Time & Condition (AGI vs CSG)			
	T2-T4		T3-T4	
	F	p<	F	p<
Total T-Score	1.247	ns (0.291)	2.322	ns (0.133)
Internal T-Score	4.433	0.05 (0.014)	7.785	0.01 (0.007)
External T-Score	0.072	ns (0.930)	0.000	ns (0.986)

Table 4d Analysis (Effect Size)

	Effect Size			
	T2-T4		T3-T4	
	AGI	CSG	AGI	CSG
Total T-Score	-0.07	0.13	-0.4	-0.09
Internal T-Score	0.13	0.28	-0.29	0.33
External T-Score	-0.15	-0.07	-0.28	-0.3

Reporting Results for Hypothesis 4d:

Total Problem Scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported between the groups over the follow-up period, with mean scores T2-T4 for AGI = 53.68 – 54.39 and for CSG = 54.63- 53.28.

Effect size analysis T2-T4 reports similar trivial scores for each group. AGI=-0.07, CSG=0.13.

Internalising Problem Scores

For groups compared to each other (AGI vs CSG)

There is statistically significant difference reported between the groups over time period, with mean scores T2-T4 for AGI = 53.75 – 52.29 and for CSG = 53.28- 50.56.

Effect size analysis T2-T4 favours the CSG (AGI is trivial =0.13, CSG is small =0.28).

Externalising Problem Scores

For groups compared to each other (AGI vs CSG)

There is no statistically significant difference reported between the groups over the follow-up period, with both groups deteriorating mean scores T2-T4 for AGI = 51.89 – 53.36 and for CSG = 53.31 – 54.03.

Effect size analysis T2-T4 reports similar trivial scores for each group. AGI is trivial =- 0.15, CSG is trivial =-0.07.

Summary:

There is evidence from the Internalising scores to show there to be a statistically significant difference between the groups at T4 favouring the CSG. The external score gives no evidence of difference between the groups. The evidence is split, but the difference is not in the direction predicted. The experimental hypothesis is rejected and the null hypothesis accepted.

"According to parent report (CBC), the AGI group will show no greater benefits on behavioural and emotional measures than the CSG group, measured prior to intervention and at one year."

Interpretation of Data and Discussion of Hypothesis 4d:

The total problem scores show the parents judging their children to have lost the gains they made over the intervention period. The same trend is present in both the AGI and CSG. This is unsurprising but disappointing. The intervention was offered in the school environment and it is to be expected that any gains made will be more easily perceived in school. The magnitude of swing back to the starting scores and even a little beyond (although this is not significant) is salutary, and plausible explanations would include the fact that parents children more difficult to manage as they get stuck into the teenage years, something that would produce a downward trend in data. Interestingly, the teachers do not report a similar global decline in the behaviours of research population, seeing instead scores remaining constant, something that can perhaps be attributed to the socialising effect of the school environment, the coping strategies of individual staff and the coherent sanction system for behaviours straying outside agreed boundaries. Attrition is a factor in this and will be discussed in Chapter 5. There is no difference between the groups on the total scores. It affects both groups equally.

The surprising finding from these data is the significant difference favouring the CSG on the Internalising score. Children are less likely to show emotional problems at home if they have received the CSG condition at a year following the cessation of intervention. This result is statistically significant, however, effect size analysis show the difference in magnitude of change in each group to be of the same order. The pattern of change measured on the Internalising scale reflects that found in Hypothesis 3 when the groups were compared over the intervention period. The parents report significant improvement for the AGI group over the intervention period that does not hold over follow-up. The CSG group on the other hand, show no change over intervention, but a steady decline over follow-up. The analyses allow this to be characterised as more children showing a slight effect (significance high, but magnitude of change low). The most ready explanation of the AGI pattern is a disappointment on behalf of the parents that more did not result from the additional input from school. The CSG pattern does tie in with the finding from the MSCS family sub-scale that reports self-concept improvements in relation to family, but an increased improvement over time is surprising.

Attrition:

The numbers of parent-completed questionnaires for children at all time points are 28 for the AGI group and 32 for the CSG group. This gives a total of 60, which is 46% of the research sample. There does not appear to be an effect of attrition that means one group is unfairly represented.

Introduction to Hypotheses 5 – 9

These hypotheses use the non-questionnaire data to examine the extent and nature of change in the same ways as previous hypotheses: considering both groups together (AGI+CSG), a non-intervention waiting-list period (T1-T2) is compared to the intervention period (T2-T3) in Hypothesis 5 and 5(sub). Still considering both groups together, change over the intervention period is compared to a follow-up period of 1 year (T3-T4) in Hypothesis 6. Hypothesis 7 compares the gains of each intervention group (AGI vs CSG) over the intervention period (T2-T3), while Hypothesis 8 compares the gains of each intervention group over the follow-up period (T3-T4).

These hypotheses relate to the findings from three sets of data

- 1) Classroom observation data 1: A protocol was developed to record time spent by the intervention groups (AGI+CSG) 'on-task' or 'off-task' in classroom activities. This observation is reported in the form of a percentage, and refers to the percentage of behaviour observed to be on-task. (possibilities include on-task/off-task/disruptive)
- 2) Classroom observation data 2: a similar protocol was developed to record the off-task behaviours of the group compared to similar behaviours from the rest of the class. This observation is a ratio of non-aggressive problem behaviours observed in the group sample versus non-aggressive problem behaviours observed in the rest of the class. The ratio of observed non-aggressive problem behaviours Group: Rest of Class is expressed as an integer value derived by dividing the average value of observed non-aggressive problem behaviours for the intervention groups (AGI & CSG for hypothesis 5 & 6 and AGI or CSG for hypotheses 7 & 8) by the average value of observed non-aggressive problem

behaviours for the rest of the class. When both values are equal the integer value is 1. When varying between 0 and 1 the observed behaviours of the group are less than those observed in the rest of the class. When varying between 1 and infinity the observed behaviours of the intervention groups are greater than those observed in the rest of the class. A value of 0.5 will indicate that average observed behaviours in the intervention groups were half those observed in the rest of the class, and a value of 2 will indicate that the observed behaviours in the intervention groups are double those observed in class.

- 3) Attendance record data: the school register of attendance was used to gather data on participant attendance, rest of class attendance and average school attendance.

Section 5.5: Hypothesis 5

Question 5. Does working in a group have an effect on a child's classroom performance and school attendance?

Do children gain in 'real-world' skills when they are involved in group interventions?

Question 5(sub). Are any effects on a child's classroom performance and school attendance reported during the group intervention greater than those over a similar period when no intervention is taking place?

If you compare children when they have been working in a group with the same children over a period when they have not, will you notice a difference in their 'real-world' skills?

What is the effect of working in a small group compared to the usual classroom experience? In order to answer this question the AGI group and the CSG group are considered together over the whole period T1-T3, from the beginning of the waiting list period to the end of intervention. It is predicted that positive feedback, group purpose and forces of group identity will have an effect of reducing 'real-world' problem behaviours observable as on- and off- task behaviours and attendance rates. Hypothesis No.5 therefore predicts that these interventions will have positive effects over whole period (T1 – T3). As this period contains a waiting-list period (T1 – T2) during which

there was no intervention, the sub hypothesis is that the gains during this T1 – T2 period will be less than the T2 – T3 intervention period.

Hypothesis No.5:

“According to independent observation and school attendance records, the intervention group (AGI+CSG) shows gains in classroom behaviour and school attendance reported from initial assessment to post-intervention (T1 – T3).”

Hypothesis No.5(sub):

“According to independent observation and school attendance records, improvements in the intervention group (AGI+CSG) over the no-intervention period (T1 – T2) will be less than over the subsequent intervention period (T2 – T3).”

The hypothesis is split into three parts (a – c) each relating to a different dataset as explained above. The full statement of Hypothesis 5 and 5(sub) is therefore as follows:

- a) According to independent observation, the intervention group (AGI+CSG) increases on-task classroom behaviour reported from initial assessment to post-intervention (T1 – T3).
- b) According to independent observation, the intervention group (AGI+CSG) shows a reduction in off-task behaviours when compared to the rest of the class, as reported from initial assessment to post-intervention (T1 – T3).
- c) According to school attendance records, the intervention group (AGI+CSG) shows increased attendance when reported from initial assessment to post-intervention (T1 – T3).

5(sub):

- a) According to independent observation, the intervention group (AGI+CSG) on-task classroom behaviour gains over the intervention period (T2-T3) will be greater than those made in the no-intervention period (T1-T2).
- b) According to independent observation, intervention group (AGI+CSG) shows a greater reduction in off-task classroom behaviour compared to the rest of the

class over the intervention period (T2-T3) than those made in the no-intervention period (T1-T2).

- c) According to school attendance records, the intervention group (AGI+CSG) shows a greater increase in attendance over the intervention period (T2-T3) than in the no-intervention period (T1-T2).

Mean scores are reported at three time points (T1: baseline, T2: pre-intervention and T3: post-intervention). ANOVA scores and effect sizes are calculated for the difference in scores for three time periods,T1-T3: the period of waiting until the end of intervention, T1-T2: the period of the waiting list control and T2-T3: the period of intervention.

Findings for Parts a-c of Hypothesis 5

Table 5 Descriptives

	t	Mean	SD	N	df
C-ob: a) On-task	1	0.72	0.15	16	2
	2	0.76	0.14	16	
	3	0.86	0.11	16	
C-ob: b) Off task ratio	1	1.02	0.74	16	2
	2	0.94	0.52	16	
	3	1.12	0.89	16	
Attendance	1	92.01	7.29	84	2
	2	93.23	8.06	84	
	3	90.29	10.70	84	

Table 5 Analysis

	Repeated Measures ANOVA & Effect Size (ES)								
	T1-T3			T1-T2			T2-T3		
	F	p<	ES	F	p<	ES	F	p<	ES
c-ob: a)	7.82	0.01 (0.002)	1.06	0.71	ns (0.413)	0.28	17.18	0.01 (0.001)	0.77
c-ob:b)	0.42	ns (0.663)	0.12	0.19	ns (0.673)	0.12	1.19	ns (0.293)	0.24
Attend ance	3.81	0.05 (0.024)	0.19	1.66	ns (0.202)	0.16	6.48	0.05 (0.013)	0.31

Hypotheses 5a & 5(sub)a

Here we report on findings from Classroom Observation a): the on-task behaviours of both the AGI and CSG groups considered as one population over the period of the waiting list (during which no intervention took place) and intervention period, and then compares the gains made during the waiting list and intervention periods.

Experimental Hypothesis 5a

“According to independent observation working in small groups increases on-task classroom behaviour reported from initial assessment to post-intervention (T1 – T3).”

Null Hypothesis 5a:

“According to independent observation working in small groups has no effect on on-task classroom behaviour reported from initial assessment to post-intervention (T1 – T3).”

Experimental Hypothesis 5(sub)a:

“According to independent observation, on-task classroom behaviour gains of the entire participant group (AGI+CSG) over the intervention period (T2-T3) will be greater than those made in the no-intervention period (T1-T2).”

Null Hypothesis 5(sub)a:

“According to independent observation, changes in on-task classroom behaviour of the entire participant group (AGI+CSG) over the intervention period (T2-T3) will be no different to those made in the no-intervention period (T1-T2).”

Reporting Results for Hypothesis 5a and 5(sub)a:

For the full period: T1-T3

Observation a): the total mean percentage scores increase from 0.72 to 0.86, indicating a substantial improvement in observed on-task behaviours taking place during T1 – T3. The p-value of 0.002 shows that this improvement is highly significant, and the effect size is moderate at 1.06.

For the waiting-list period: T1-T2

Observation a): the total mean percentage scores increase from 0.72 to 0.76, a non-significant improvement in observed on-task behaviours ($p=0.413$). The effect size is small at 0.28.

For the intervention period: T2-T3

Observation a): the total mean percentage scores increase from 0.76 to 0.86, indicating a substantial improvement in observed on-task behaviours taking place during T2 – T3. This change is highly significant ($p=0.001$), and the effect size is moderate at 0.77.

Summary:

These data give compelling evidence to accept that on-task behaviour increases during the intervention period compared to no change over the waiting list period. The evidence reported above gives good reason to reject the null hypothesis in both cases, allowing experimental hypothesis 5a and 5(sub)a to be re-stated

Experimental Hypothesis 5a:

“According to independent observation working in small groups increases on-task classroom behaviour reported from initial assessment to post-intervention (T1 – T3).”

Experimental Hypothesis 5(sub)a:

“According to independent observation, on-task classroom behaviour gains of the entire participant group (AGI+CSG) over the intervention period (T2-T3) will be greater than those made in the no-intervention period (T1-T2).”

Interpretation of Data and Discussion of Hypothesis 5a and 5(sub)a:

The on-task behaviours of the intervention group considered as one population behave as predicted in the hypotheses. From initial assessment to assessment just prior to intervention, there is no change in the scores and the magnitude of change as measured by the effect size is small. During intervention however these scores increase dramatically with a moderate size meaning that the average child has moved from the 50th to the 80th centile over the time period in question. It is with reasonable confidence that one can say that the groupwork intervention has an effect in the classroom, making children more able to stay on-task during classroom activities. It is reasonable to suggest that the positive attention given to these children during the group sessions makes them more able to invest value in their work, making them less likely to be distracted. The numbers are small, but this makes the highly significant change more persuasive over the intervention period.

Attrition:

The problem here is not with attrition, it is with low participant numbers. 16 from a cohort of over 100 is a small sample and the explanation for this lies in the complexity and labour intensiveness of the observation programme. This meant that only the later groups of children were available for inclusion in the waiting list observations at time 1. These children were the only ones admissible for inclusion in the other time points.

Hypothesis 5b & 5(sub)b

This report is on findings from Classroom Observation b): the off-task behaviours of the AGI+CSG groups compared to the rest of the class

Experimental Hypothesis 5b:

“According to independent observation, working in small groups reduces off-task behaviours when compared to the rest of the class, as reported from initial assessment to post-intervention (T1 – T3).”

Null Hypothesis 5b:

“According to independent observation, working in small groups has no effect on off-task behaviours when compared to the rest of the class, as reported from initial assessment to post-intervention (T1 – T3).”

Experimental Hypothesis 5(sub)b:

“According to independent observation, on-task classroom behaviour gains of the (AGI+CSG) group expressed as a ratio of total class on-task behaviours will be greater over the intervention period (T2-T3) than those made in the no-intervention period (T1-T2).”

Null Hypothesis 5(sub)b:

“According to independent observation, changes in on-task classroom behaviour gains of the (AGI+CSG) group expressed as a ratio of total class on-task behaviours over the intervention period (T2-T3) will be no different to those made in the no-intervention period (T1-T2).”

Reporting Results for Hypothesis 5b and 5(sub)b:

For the full period: T1-T3

Observation b): the total mean ratio scores for off-task behaviours increase slightly T1-T3 from 1.02 to 1.12, indicating that off-task behaviours of the intervention group are increasing relative to the rest of the class. This increase is a long way from being significant, ($p=0.663$) and the effect size trivial at 0.12

For the waiting-list period: T1-T2

Observation b): the total mean ratio scores for off-task behaviours decrease slightly T1-T2 from 1.02 to 0.94, this is not significant, ($p=0.673$) and the effect size trivial at 0.12

For the intervention period: T2-T3

Observation b): the total mean ratio scores for off-task behaviours increase slightly T2-T3 from 0.94 to 1.12, indicating that off-task behaviours of the intervention group are increasing relative to the rest of the class. This increase is a long way from being significant, ($p=0.293$), and the effect size small at 0.24

Summary:

There is no significant change during any of the time periods and the magnitude of change is small in both cases. On the evidence presented above, the experimental hypotheses 5b and 5(sub)b are both rejected and the null hypothesis accepted in each case:

Null Hypothesis 5b

"According to independent observation, working in small groups does nothing to reduce off-task behaviours when compared to the rest of the class, as reported from initial assessment to post-intervention (T1 – T3).

Null Hypothesis 5(sub)b

"According to independent observation, on-task classroom behaviour gains of the (AGI+CSG) group expressed as a ratio of total class on-task behaviours will be no different over the intervention period (T2-T3) from those made in the no-intervention period (T1-T2)."

Interpretation of Data and Discussion of Hypothesis 5b:

These data relate the intervention group and their off-task behaviours to the rest of the class. On this evidence there is no difference between them. This could mean that the whole class was experiencing a tighter teaching method during the T2-T3 period, this would allow the gains on on-task behaviours reported in hypothesis 5a and 5(sub)a to be accounted for in generalised classroom effects of increased attention. It is worth noting that the off-task behaviours of children screened for behavioural issues are reported to be at the same level (ratio of 1.02) as the rest of the class before any intervention has taken place. This is counter-intuitive, and it could be that the low numbers have allowed localised conditions of school and selection to provide unrepresentative data.

Hypotheses 5c & 5(sub)c

This report is on findings from the school records: the attendance of both the AGI and CSG groups is expressed as a percentage

Experimental Hypothesis 5c:

“According to school attendance records, working in a small group based on mutual regard increases attendance when reported from initial assessment to post-intervention (T1 – T3).”

Null Hypothesis 5c:

“According to school attendance records, working in a small group based on mutual regard has no positive effect on attendance when reported from initial assessment to post-intervention (T1 – T3).”

Experimental Hypothesis 5(sub)c:

“According to school attendance records, working in a small group based on mutual regard increases attendance more over the intervention period (T2-T3) than in the no-intervention period (T1-T2).”

Null Hypothesis 5(sub)c:

“According to school attendance records, there is no positive difference in school attendance over the intervention period (T2-T3) compared to the no-intervention period (T1-T2) while working in a small group based on mutual regard.”

Reporting Results for Hypothesis 5b and 5(sub)b:

For the full period: T1-T3

Attendance scores: percentage attendance falls T1-T3 (92.01 – 90.29), this is significant at $p < 0.05$ ($p = 0.024$). The effect size is trivial at 0.19.

For the waiting-list period: T1-T2

Attendance scores: percentage attendance rises T1-T2: (92.01 – 93.23), this is not significant ($p = 0.202$). The effect size of 0.16 is trivial.

For the intervention period: T2-T3

Attendance scores: percentage attendance falls T2-T3 (93.23 – 90.29), this is significant at $p < 0.05$ ($p = 0.013$). The effect size of 0.31 is small.

The evidence presented above is not sufficient for the null hypothesis to be rejected and is re-stated for hypotheses 5c and 5(sub)c is re-stated below:

Summary:

Null Hypothesis 5c:

“According to school attendance records, working in a small group based on mutual regard has no positive effect on attendance when reported from initial assessment to post intervention (T1 – T3).”

Null Hypothesis 5(sub)c:

“According to school attendance records, there is no positive difference in school attendance over the intervention period (T2-T3) compared to the no-intervention period (T1-T2) while working in a small group based on mutual regard.”

Interpretation of Data and Discussion of Hypothesis 5c and 5(sub)c:

The lack of significant change over the waiting list period is in keeping with the predictions made in the hypothesis. With no intervention there is no change in attendance. There is significant change over the intervention period, but it is a decrease in attendance that produces the change. The numbers of 84 are quite respectable and

there can be no claim of variable data due to low numbers. The effect size shows that the magnitude of change to be small at 0.31, but it is not far away from the level judged to be noteworthy in educational interventions. The change is in the opposite direction to that predicted in that while the intervention groupwork sessions are taking place, the children involved in the study become less likely to attend school. There does not seem to be any good reason for this, since the extra attention that the groupwork affords should allow the children to feel more valued in school and pleased to be part of something unusual in school. Certainly the data from the questionnaires over the same period reveal a rise in positive behaviour that was attributed to a 'selection effect' of anticipation. It is possible that the children have such a positive experience in the group sessions that they realise how sub-standard the usual curriculum delivery is and decide to vote with their feet. Practitioners have noted that a feature of groups run for older children is that participants frequently remark that the group sessions are the one thing they make sure they are in school for, but this does not account for the significant decrease in attendance that includes the primary schoolchildren. Perhaps this finding will be illuminated by subsequent analysis, and perhaps it will be revealed as an anomaly.

Attrition:

The number of participants of 84 from a possible 130 shows a high level of incomplete data, until the nature of the waiting-list methodology is appreciated. The first cohort of children received their interventions without receiving a waiting list data collection, this was due to practical necessity of the type frequently encountered in running programmes of delivery alongside evaluations. The relatively low numbers of participant data (64% of the research population) can be accounted for in this way.

Section 5.6: Hypothesis 6

Question 6. Do the effects on a child's classroom performance and school attendance made while the groups are running last after the group is finished?

If you survey those involved in groupwork at intervals after the interventions are over will you still be able to detect any impact made in 'real-world' behaviours?

Are the 'real-world' effects of working in a group sustained after intervention has finished? In order to answer this question both AGI and CSG groups are considered together and time points T2, T3 and T4 are considered. The input common to both AGI and CSG groups is intended to have a lasting impact in the way a child behaves and based on the reading of the literature in Chapter One, this effect in 'real-world' settings should be present one year after intervention. It is therefore predicted that any improvement for the intervention groups over the intervention period T2-T3 will be sustained one year later at T4. The hypothesis looks for significant improvement over intervention and no significant change over the follow-up period.

Hypothesis No.6:

"According to independent observation and school attendance records, gains made by the intervention group (AGI+CSG) in classroom behaviour and school attendance over the intervention period (T2-T3) are sustained over the follow-up period (T3-T4)."

The hypothesis is split into three parts (a – c) each relating to a different dataset as explained above. The full statement of Hypothesis 6 is therefore as follows:

- a) According to independent observation, gains made by the intervention group (AGI+CSG) in on-task classroom behaviour over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).
- b) According to independent observation of the intervention group (AGI+CSG), (AGI+CSG):[rest of class] off-task behaviour ratio reduces over the intervention period (T2 – T3) and is sustained over the follow-up period (T3 – T4).
- c) According to school attendance records, the gains by the intervention group (AGI+CSG) in percentage attendance over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4).

Mean scores are reported at three time points (T2: pre-intervention, T3: post-intervention and T4 1-year post intervention). ANOVA scores and Effect sizes are calculated for the difference in scores for three time periods, T2-T4: the period of intervention and follow-up, T2-T3: period of Intervention and T3-T4 the period of follow-up one year post intervention. For the purposes of discriminating between the hypothesis and the null hypothesis it is the comparison of these latter two time periods that will be examined. The population of AGI and CSG groups are considered here as a whole. Hypothesis 6 is concerned with the effects of the total groupwork population over a year follow-up.

Table 6 Descriptives

	t	Mean	SD	N	df
Observation a) On-task	2	0.81	0.12	43	2
	3	0.81	0.16	43	
	4	0.77	0.19	43	
Observation b) Off-task ratio	2	1.28	0.94	31	2
	3	1.05	0.54	31	
	4	2.47	3.09	31	
Attendance	2	93.00	7.10	79	2
	3	92.18	10.50	79	
	4	89.20	14.41	79	

Table 6 Analysis

	Repeated Measures ANOVA & Effect Size (ES)								
	T2-T4			T2-T3			T3-T4		
	F	p<	ES	F	p<	ES	F	p<	ES
c-ob: a)	1.00	ns (0.371)	0.00	0.00	ns (0.985)	0.00	1.17	ns (0.287)	0.23
c-ob: b)	5.76	0.01 (0.005)	-0.51	3.31	ns (0.079)	0.30	6.50	0.05 (0.016)	-0.63
Attend ance	4.80	0.01 (0.009)	0.33	0.52	ns (0.474)	0.09	5.86	0.05 (0.018)	0.24

Hypothesis 6a

This part of the Hypothesis reports on findings from Classroom Observation a): the on-task behaviours of both the AGI and CSG groups.

Experimental Hypothesis 6a

"According to independent observation, gains made in on-task classroom behaviour over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4)."

Null Hypothesis 6a

"According to independent observation, there is no difference between any changes in on-task classroom behaviour over the intervention period (T2 – T3) and changes in the follow-up period (T3 – T4)."

Reporting Results for Hypothesis 6a:

For the full period: T2-T4

Observation a): the total mean percentage scores decrease from 0.81 to 0.77, indicating a deterioration in observed on-task behaviours taking place during T2 – T4. This change is not significant ($p=0.371$), and the effect size is small at 0.25.

For the intervention period T2-T3

Observation a): the total mean percentage scores show no change (0.81 at T2 0.81 at T3), this is not significant ($p=0.985$). The effect size is trivial at 0.00.

For the follow-up period T3-T4

Observation a): the total mean percentage scores decrease from 0.81 to 0.77, indicating a deterioration in observed on-task behaviours taking place during T3 – T4. This change is not significant ($p=0.287$), and the effect size is small at 0.23.

Summary:

Hypothesis 6a predicts gains over the intervention period T2-T3 to be sustained over the follow-up period T3-T4. There is no significant deterioration of observed on task behaviours T3-T4, there is no significant gain T2-T3. The emphasis is on the follow-up period and the magnitude of change over this period is small, allowing for an interpretation of follow-up stability. However the weight of evidence compels the acceptance of the null hypothesis 6a:

Null Hypothesis 6a:

According to independent observation, there is no difference between any changes in on-task classroom behaviour over the intervention period (T2 – T3) and changes in the follow-up period (T3 – T4).

Interpretation of data and discussion of hypothesis 6a:

When both intervention groups are considered together (AGI and CSG) the percentage of behaviours observed to be on-task is 81% over the intervention period T2-T3. This seems high, but perhaps reflects the disproportionate amount of attention that off-task behaviours demand, and the behaviours that children will display before being labelled as disaffected. The null hypothesis is accepted on the basis that although this high level of observed on-task behaviours is sustained at follow-up of 1 year, and the percentage fall to 77% is not registered as significant, this pattern is at variance with the pattern in hypothesis 5a, which reported a significant improvement in the on-task scores over the T2-T3 period.

Attrition:

The reduced numbers of 43 is a third of the research population and reflects the methodological difficulties of conducting observations on children dispersed between classes in a high school. Whereas in the primary setting, the intervention children would have perhaps formed half of the class population and observations at any one time point could be completed in four sessions (2am and 3pm), in the secondary school, the 16 children might be split between 3 high schools, each with 8 classes. Under these circumstances with limited funds available, children missed the observations at a rate over and above that of natural attrition.

Hypothesis 6b

This part of the Hypothesis reports on findings from Classroom Observation b): the off-task behaviours of the AGI+CSG groups compared to the rest of the class

Experimental Hypothesis 6b

“According to independent observation, intervention group: rest of class off-task behaviour ratio reduces over the intervention period (T2 – T3) and is sustained over the follow-up period (T3 – T4).”

Null Hypothesis 6b

“According to independent observation, there is no difference between changes to the intervention group: rest of class off-task behaviour ratio over the intervention period (T2 – T3) and changes over the follow-up period (T3 – T4).”

Reporting Results for Hypothesis 6b:

For the full period: T2-T4

Observation b): the total mean ratio scores for off-task behaviours increases T2-T4 from 1.28 to 2.47, indicating that off-task behaviours of the intervention group are increasing relative to the rest of the class. This increase is highly significant, ($p=0.005$) and the effect size of 0.51 is small.

For the intervention period T2-T3

Observation b): the total mean ratio scores for off-task behaviours decrease slightly T2-T3 from 1.28 to 1.05, this is not significant, ($p=0.079$) and the effect size of 0.30 is small.

For the follow-up period T3-T4

Observation b): the total mean ratio scores for off-task behaviours increase T3-T4 from 1.05 to 2.47, indicating that off-task behaviours of the intervention group are increasing relative to the rest of the class. This increase is significant at $p<0.05$, ($p=0.018$), and the effect size of 0.63 is moderate.

Summary:

Hypothesis 6b predicts improvements in observed off-task behaviours relative to the rest of the class. This does look like it is happening T2-T3, but in nothing like the amounts to produce significant change. In the period T3-T4 the reverse is true: the experimental group off-task behaviours increase relative to the rest of the class and show a moderate magnitude of change. On this evidence the experimental Hypothesis 6b is rejected and the null hypothesis accepted:

Null Hypothesis 6b:

According to independent observation, any improvement in intervention group:[rest of class] off-task behaviour ratio over the intervention period (T2 – T3) will not be sustained over the follow-up period (T3 – T4).

Interpretation of data and discussion of hypothesis 6b:

The values of observed non-aggressive problem behaviours show little variability over the intervention period T2-T3. The values of 1.28 and 1.05 show behaviours on a par with those observed in the rest of the class, and of the same order as those observed during the waiting list period, T1-T2, the greater numbers ($n=31$) adding validity to the findings compared to Hypothesis 5 where $n=16$. The rather bland rejection of the experimental hypothesis and acceptance of the null hypothesis that any improvement is not sustained at follow-up does not do justice to the extraordinary and significant increase of observed problem behaviours. These are observed to be twice as frequent as those observed in the rest of the class, and the difference between T3 and T4 is significant at $p<0.05$. In fact, what looks like a marginal improvement T2-T3 is close to significant ($p=0.07$) with a moderate effect size supporting a notable magnitude of change, making the subsequent decline T3-T4 more pronounced. Taken at face value the explanations for this turn-around are hard to find. The primary intervention groups have all changed school and will be mixed in with a higher proportion at non-intervention children whom might reasonably be expected to display higher levels of problem behaviours, leaving the face-value explanation to be that an unanticipated process is active as a result of the interventions, and this has made the research participants more likely to be disruptive in class. Deeper explanations are harder to find, but might possibly rest upon a 'self-confidence effect', in that self confidence instilled over the intervention period, but not registering in the 'class of origin' makes their behaviour less manageable a year later. This is a tenuous explanation and the more persuasive argument is the methodological one. In support of this explanation it is worth noting that over the same period observations conducted by the same observers on the same children (although there is a reduction in participant numbers from 43-31) observe on task behaviours to have not differed significantly. The non-significant decrease in on-task behaviours is at some odds with the massive increase in observed problem behaviours. While it is true that an increase in observed problem behaviours might be expected, the extraordinary magnitude of increase is persuasively explained by the difficulties in conducting the off-task observation in a classroom containing 2 or three

subjects without having a bias towards more detailed observation of their behaviours. If only two children from a class of 30 are identified prior to the observation starting, it could well be difficult to maintain the impartiality in the observer's gaze. Although this extraneous variability affects both AGI and CSG intervention groups equally, it would be most noticeable when the whole intervention group was being compared to the rest of the class, because the children identified for observation in this argument receive disproportionate attention from the observer when they are only one or two in any one class.

Hypothesis 6c

This report is on findings from the school records: the attendance of both the AGI and CSG groups is expressed as a percentage

Experimental Hypothesis 6c

"According to school attendance records, the gains in percentage attendance over the intervention period (T2 – T3) are sustained over the follow-up period (T3 – T4)."

Null Hypothesis 6c

"According to school attendance records, there is no difference between changes in percentage attendance over the intervention period (T2 – T3) and changes over the follow-up period (T3 – T4)."

Reporting Results for Hypothesis 6c

For the full period: T2-T4

Attendance scores: percentage attendance falls T2-T4 (93.00 – 89.20). This is significant at $p < 0.01$ ($p = 0.009$). The effect size of 0.33 is small.

For the intervention period T2-T3

Attendance scores: percentage attendance falls T2-T3: (93.00 – 92.18). This is not significant ($p = 0.474$). The effect size of 0.09 is trivial.

For the follow-up period T3-T4

Attendance scores: percentage attendance falls T3-T4 (92.18- 89.20). This is significant at $p < 0.05$ ($p = 0.018$). The effect size of 0.24 is small.

Summary:

Hypothesis 6c predicts improvements in school attendance sustained over follow-up. The evidence is that there is no significant change in attendance T2-T3, but a sharp and significant decline T3-T4. On this evidence, the experimental hypothesis 6c is rejected and the null hypothesis accepted.

Null Hypothesis 6c:

According to school attendance records, any gains in percentage attendance over the intervention period (T2 – T3) are not sustained over the follow-up period (T3 – T4).

Interpretation of data and discussions of hypothesis 6c:

Attendance holds to the same order of 93% over the intervention period T2-T3. This is followed by a sharp decline to 89% that is significant at $p < 0.05$ ($p = 0.018$). At face value this looks as if an effect of intervention is a reduced inclination to attend school.

Furthermore these data seem to support those of hypothesis 6b that report an increase in problem behaviours in the classroom, an indicator of engagement with school. The process underlying the fall in 'face-value' explanations of decreased attendance is equally hard to find as that underlying the increased problem behaviours. It is counter-intuitive that a benign intervention sensitively delivered could have such a demonstrable negative effect. It is the change of school explanation that is more appealing, if only through the lens of Occam's Razor (Occam's Razor states that the simplest explanation is likely to be the right one). There are acknowledged difficulties in maintaining attendance levels between primary and secondary school with performance tables for the period 1997 showing that in North Tyneside the average attendance for the primary schools in the study was 95% while those for John Spence and Norham High Schools (the schools the majority of pupils in the study transferred to) were 91% and 87% respectively (source: <http://www.dfes.gov.uk/cgi-bin/shleap>). This information gives credence to explanation looking to the high percentage of children in the study changing school over the follow-up decline. The lack of at-risk control over this period makes this explanation factually inconclusive, however it is persuasive.

Section 5.7: Hypothesis 7

Question 7. Do the effects on a child's classroom performance and school attendance depend on what goes on in the group?

If you compare a group who are engaged in child-focussed developmental and therapeutic activities (AGI) with a group engaged in curriculum tasks (CGI), will you notice a difference in 'real-world' skills?

Is there any difference between the groups in the short-term? To answer this the impact on the AGI group can be compared to the impact on the CSG group the intervention period T2-T3. The groups are experimentally controlled for variables such as time of day, duration, facilitator and general child-centred approach, but are different in the content of the sessions delivered. Because the AGI sessions are more clearly focussed on the enrichment of relationships and the group encouraged to consider positive management of situations outside the group sessions, it is predicted that the AGI group will be more effective in reducing problem behaviours and increasing self-esteem.

Hypothesis 7:

According to independent observation and school attendance records, the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.

The Hypothesis is split into three parts (a – c) each relating to a different dataset, therefore the full statement of Hypothesis 7 is as follows:

- a) According to independent observation over the intervention period, the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.**
- b) According to independent observation over the intervention period, the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.**

- c) According to school attendance records over the intervention period, the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.

The analyses applied are the repeated measures ANOVA and Effect Size to measure the magnitude of change.

Findings for Hypothesis 7

Table 7 Descriptives

	t	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Observation a)	2	0.80	0.14	45	0.78	0.15	43	1
	3	0.82	0.13	45	0.81	0.23	43	
Observation b)	2	1.45	0.99	36	1.41	0.88	35	1
	3	1.44	1.08	36	1.50	1.22	35	
Attendance	2	92.95	8.66	52	93.10	7.42	51	1
	3	90.85	11.08	52	90.57	10.32	51	

Table 7 Analysis

	Repeated measures ANOVA		Effect Size (ES)	
	Time by Condition (AGI vs CSG)		T2-T3	
	F	p<	AGI	CSG
Observation a)	0.02	ns (0.897)	0.15	0.15
Observation b)	0.24	ns (0.629)	0.01	0.08
Attendance	0.04	ns (0.841)	0.21	0.28

Hypothesis 7a

This part of the Hypothesis reports on findings from Classroom Observation a): the on-task behaviours of both the AGI and CSG groups.

Experimental Hypothesis 7a

“According to independent observation over the intervention period, the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Null Hypothesis 7a

“According to independent observation over the intervention period, the groupskills intervention (AGI) is no more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Reporting Results for Hypothesis 7a:

For groups compared to each other (AGI vs CSG)

Mean scores for the AGI group rise from 0.80 – 0.82 with a small effect size of 0.15, as do those for the CSG group, beginning from a lower score (0.78- 0.81) with an effect size 0.15. The difference between the groups is not significant ($p=0.897$) and the combined effect size is trivial at 0.00

Summary:

The lack of any significant difference between the AGI and CSG groups at post intervention mean that the experimental hypothesis 7a is rejected and the null hypotheses accepted:

Null Hypothesis 7a:

“According to independent observation over the intervention period, the groupskills intervention (AGI) is no more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Interpretation of Data and Discussion of Hypothesis 7a:

When both groups are considered together (AGI+CSG) over the intervention period T2-T3 there is no significant change in observed on-task behaviours. This finding is supported by hypothesis 6a that measured change over the same period, but is contrary to the findings of Hypothesis 5a that reported significant increase in on-task behaviours over the same period. Actually there is an increase reported by this hypothesis that looks

as if it might be significant from 79% to 82%, but this will remain a trend in the data that is either corroborated by 5a or undermined by 6a. When the two groups are compared there is nothing to choose between them, giving an unambiguous finding that the content of the group sessions appears to have no differential effect on observed on-task behaviours.

Hypothesis 7b

This part of the Hypothesis reports on findings from Classroom Observation b): the off-task behaviours of the AGI+CSG groups compared to the rest of the class

Experimental Hypothesis 7b

“According to independent observation over the intervention period, the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Null Hypothesis 7b

“According to independent observation over the intervention period, the groupskills intervention (AGI) is no more effective than the small-group curriculum studies Intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Reporting Results for Hypothesis 7b:

For groups compared to each other (AGI vs CSG)

The AGI scores fall slightly (1.45-1.44) with a tiny effect size of 0.01, while CSG scores rise (1.41- 1.50) with a similarly tiny effect size of 0.08. There is no significant difference between the groups ($p=0.629$) and the combined effect size is negative, but tiny at -0.09

Summary:

The lack of any significant difference between the AGI and CSG groups at post intervention mean that the experimental hypothesis 7b is rejected and the null hypotheses accepted:

Null Hypothesis 7b:

“According to independent observation over the intervention period, the groupskills intervention (AGI) is no more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Interpretation of Data and Discussion of Hypothesis 7b:

When both groups are considered together (AGI+CSG) over the intervention period T2-T3 there is no significant change in observed on-task behaviours. This finding is supported by both Hypothesis 6b that measured change over the same period, and the findings of Hypothesis 5b. The trends in the non-significant values for 7b are barely present and the difference between the groups barely noticeable even before analysis of variance reveals a non-significant p-value. When the two groups are compared there is nothing to choose between them, giving an unambiguous finding that the content of the group sessions appears to have no effect on observed non-aggressive off-task behaviours.

Hypothesis 7c

This part of the Hypothesis reports on findings from the school records: the attendance of both the AGI and CSG groups is expressed as a percentage

Experimental Hypothesis 7c

“According to school attendance records over the intervention period, the groupskills intervention (AGI) is more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Null Hypothesis 7c

“According to school attendance records over the intervention period, the groupskills intervention (AGI) is no more effective than the small-group curriculum studies intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Reporting Results for Hypothesis 7c:

For both groups compared to each other (AGI vs CSG)

The AGI mean scores fall 92.95-9.85 with a small effect size of 0.21, and CSG scores fall similarly 93.10-90.57 with a small effect size of 0.28. There is no significant

difference between the groups ($p=0.841$) and the combined effect size (AGI-CSG) is tiny at 0.07.

Summary:

The lack of any significant difference between the AGI and CSG groups at post Intervention mean that the experimental hypothesis 7c is rejected and the null hypotheses accepted:

Null Hypothesis 7c:

“According to school attendance records over the Intervention period, the groupskills Intervention (AGI) is no more effective than the small-group curriculum studies Intervention (CSG) in promoting gains in classroom behaviour and school attendance.”

Interpretation of Data and Discussion of Hypothesis 7c:

When both groups are considered together (AGI+CSG) over the intervention period T2-T3 there is a significant decrease in attendance levels. This finding exactly mirrors that of Hypothesis 5c (right down to the reported values) but is contrary to the findings of Hypothesis 6c that reported no change over the intervention period. The trend in all cases appears to be decreasing attendance over the intervention period. When the two groups are compared there is nothing to choose between them, giving an unambiguous finding that the content of the group sessions appears to have no effect on attendance

Attrition:

The number of participants is recorded at 103, which represents 80% of the research population with equal numbers of both intervention conditions.

Section 5.8: Hypothesis 8

Question 8. How do any differences in children's classroom performance and school attendance between the groups alter as time goes by?

If you survey the respondents at intervals after the interventions are over will you still be able to detect differences in the 'real-world' skills held by the different groups?

Is it possible to distinguish between the intervention groups at a follow-up point of one year? To answer this the AGI and CSG groups are compared at one year following intervention to see if there is any difference between the non-questionnaire data. The AGI intervention is considered to be more able to affect the multiple contexts in which an individual will express disaffection also the multiple areas of emotional and social function that need to be addressed in order to provide lasting change. It is therefore predicted that the positive impact on AGI intervention group will be greater at the follow-up point of one year (T4) than the CSG group whose intervention has focussed on using the group to achieve curriculum tasks.

Hypothesis 8:

According to independent observation and school attendance records, there will be a measurable difference between gains in classroom behaviour and school attendance resulting from the groupskills intervention (AGI) and those gains promoted by the curriculum studies intervention (CSG) at a follow-up point of one year.

The Hypothesis is split into three parts (a – c) each relating to a different dataset. Therefore the full statement of Hypothesis 8 is as follows:

- a) According to independent observation at follow-up, there will be a measurable difference between AGI and CSG groups, with scores made in on-task classroom behaviour resulting from the groupwork intervention (AGI) exceeding those scores promoted by the curriculum studies intervention (CSG) at the follow-up point of one year.
- b) According to independent observation at follow-up, there will be a measurable difference between AGI and CSG groups, with AGI: [rest of class] off-task behaviour ratio scores lower than CSG: [rest of class] off-task behaviour ratio scores at the follow-up point of one year.
- c) According to school attendance records at follow-up, there will be a measurable difference between AGI and CSG groups, with in percentage attendance by the AGI group exceeding percentage attendance made by the CSG group at a follow-up point of one year.

Findings for Hypothesis 8

Table 8 Descriptives

	T	AGI			CSG			df
		Mean	SD	N	Mean	SD	N	
Observation a)	2	0.81	0.12	21	0.80	0.13	22	2
	3	0.81	0.15	21	0.80	0.16	22	
	4	0.75	0.18	21	0.79	0.19	22	
Observation b)	2	1.25	1.09	16	1.31	0.80	15	2
	3	0.96	0.51	16	1.14	0.58	15	
	4	2.59	2.95	16	2.35	3.34	15	
Attendance	2	93.99	6.09	40	91.99	7.97	39	2
	3	92.32	11.09	40	92.03	10.01	39	
	4	87.72	17.92	40	90.72	9.59	39	

Table 8 Analysis (ANOVA)

	Repeated Measures ANOVA – Time & Condition (AGI vs. CSG)			
	T2-T4 df=1		T3-T4 df=1	
	F	p<	F	p<
Observation a)	0.44	ns (0.644)	0.50	ns (0.483)
Observation b)	0.11	ns (0.895)	0.14	ns (0.714)
Attendance	1.97	ns (0.143)	1.81	ns (0.182)

Table 8 Analysis (Effect Size)

	Effect Size			
	T2-T4		T3-T4	
	AGI	CSG	AGI	CSG
Observation a)	-0.38	-0.06	-0.36	-0.06
Observation b)	0.59	0.42	0.75	0.49
Attendance	0.46	0.14	0.31	0.13

Hypothesis 8a

This part of the Hypothesis reports on findings from Classroom Observation a): the on-task behaviours of both the AGI and CSG groups.

Experimental Hypothesis 8a

“According to independent observation at follow-up, there will be a measurable difference between AGI and CSG groups, with gains made in on-task classroom behaviour resulting from the group work intervention (AGI) exceeding those gains promoted by the curriculum studies intervention (CSG) at the follow-up point of one year.”

Null Hypothesis 8a

“According to independent observation, there will be no measurable difference between gains made in on-task classroom behaviour resulting from the devised groupwork intervention (AGI) and those gains promoted by the curriculum studies intervention (CSG) at a follow-up point of one year.”

Reporting Results for Hypothesis 8a:

For groups compared to each other (AGI vs. CSG)

The analyses do not give evidence for distinguishing between the AGI and CSG groups either over the period of intervention or follow-up. There is no significant difference between these groups on-task behaviour according to Independent observation at any time point. The mean scores of each group fall steadily, but broadly mirror each other in

the rate at which they fall, the only difference being in the AGI group scoring higher in initial scores (AGI=0.81, CSG=0.80) and lower in the final scores (AGI=0.75, CSG=0.79). Indicating a higher degree of change. The effect sizes bear this out with the small effect size of the AGI group at 0.36 compared to the tiny effect size of the CSG group at 0.06.

Summary:

The lack of any significant difference between the AGI and CSG groups at a follow-up point of one-year mean that the experimental hypothesis 8a is rejected and the null hypothesis accepted:

Null Hypothesis 8a:

“According to Independent observation, there will be no measurable difference between gains made in on-task classroom behaviour resulting from the devised groupwork intervention (AGI) and those gains promoted by the curriculum studies intervention (CSG) at a follow-up point of one year.”

Interpretation of data and discussion of Hypothesis 8a:

The prediction of the hypothesis is that there will be a detectable and significant difference between the groups at one-year follow-up, with the scores for the AGI group exceeding those of the CSG group. There is no significant difference between the groups at one year and no significant effect of time on the scores when both groups are considered together. Hypothesis 5 observed a significant increase in the same on-task behaviours from the beginning of the waiting-list period T1 to the end of intervention T3. This finding is relevant here because Hypothesis 8 supports Hypothesis 6 in asserting that those scores are sustained by both groups at one-year.

However, the notion that there will be a difference between the groups is not supported, and it is possible only to conclude that AGI and CSG interventions are equally effective in promoting and supporting on-task behaviours in the classroom. Deeper consideration of the processes involved allows the proposition that as the thrust of the AGI intervention is the encouragement of independence in thought and reflective action, then it could be that there is an active decision to selectively engage for the Externalising children, coupled with an effect that the introverted behaviours are less apparent due to increased

confidence, leading to an decrease in observed on-task behaviours due to increased social interaction. The scores T3-T4 are not significantly different and therefore it could be asserted that if this decrease exists due to the processes proposed, it is balanced by an corresponding increase in on-task behaviours.

Hypothesis 8b

This part of the hypothesis reports on findings from Classroom Observation b): the off-task behaviours of the AGI+CSG groups compared to the rest of the class

Experimental Hypothesis 8b

“According to independent observation at follow-up, there will be a measurable difference between AGI and CSG groups, with reductions in AGI:[rest of class] off-task behaviour ratio exceeding CSG:[rest of class] off-task behaviour ratio at the follow-up point of one year.”

Null Hypothesis 8b

“According to independent observation, there will be no measurable difference between reductions in AGI:[rest of class] off-task behaviour ratio on-task classroom behaviour and CSG:[rest of class] off-task behaviour ratio at a follow-up point of one year.”

Reporting Results for Hypothesis 8b:

For both groups compared to each other (AGI vs CSG)

There is no significant difference between the groups reported at any point. The mean scores show the AGI group with a greater rate of improvement T2-T3 (1.09 – 0.96 ES=0.33 compared to CSG 1.31 – 1.14, ES=0.24) and a greater rate of decline T3-T4 (0.96 – 2.59, ES=0.75 compared to CSG 1.14 – 2.35, ES=0.49. The effect size scores show the magnitude of change in the AGI group to be greater than the CSG group, but when combined to determine any difference between the groups the effect sizes are trivial; at 0.09 over the intervention period T2 – T3 and small at 0.26 over the follow-up period.

Summary:

The lack of any significant difference between the AGI and CSG groups at a follow-up point of one-year mean that the experimental hypothesis 8b is rejected and the null hypothesis accepted:

Null Hypothesis 8b:

“According to independent observation, there will be no measurable difference between reductions in AGI:[rest of class] off-task behaviour ratio on-task classroom behaviour and CSG:[rest of class] off-task behaviour ratio at a follow-up point of one year.”

Interpretation of data and discussion of Hypothesis 8b:

The acceptance of the null hypothesis means that the prediction that the AGI intervention will be more effective in reducing non-aggressive off-task behaviours than the CSG group is not supported by the data. There is no difference between the groups at T4 and we must conclude that the content of the group sessions is not a decisive factor in how an at-risk group behaves in the classroom. However this conclusion hides an interesting and significant finding from the data: that those involved in group work are observed to display more non-aggressive off-task behaviours than the rest of the class as time goes on. When both groups are considered together (AGI+CSG) there is a significant ($p < 0.05$) increase in the ratio of group behaviours to rest of class behaviours (1.05 at T3 to 2.47 at T4) and there is a magnitude of change with an effect size over 1.

Why would two benign interventions produce such an effect? The counter-intuitive aspect of this findings leads to a hard look at the methodology of this observation protocol, and it is true that the validity and reliability are not established. However to ignore these surprising (and pervasive: see hypothesis 6) findings is to miss an opportunity to explore the possible processes at work in child groups – on the basis that there is no smoke without fire this finding is worth attention. Where this hypothesis can add to the discussion of hypothesis 6 is in the finding that the groups AGI and CSG are equally affected by the observed increase in non-aggressive off-task behaviours. As discussed in Hypothesis 6, the learning processes that the groups were led through were very different, but the effects are similar on this measure, also although the off-task behaviours increasing at the same time as stability in observed on-task behaviours

seems a difficult tension, the rationale advanced in Hypothesis 6 makes this at the very least a possibility.

Hypothesis 8c:

This part of the Hypothesis reports on findings from the school records: the attendance of both the AGI and CSG groups is expressed as a percentage

Experimental Hypothesis 8c

"According to school attendance records at follow-up, there will be a measurable difference between AGI and CSG groups, with percentage attendance by the AGI group exceeding percentage attendance made by the CSG group at a follow-up point of one year."

Null Hypothesis 8c

"According to school attendance records, there will be no measurable difference between the percentage attendance in the AGI group and the percentage attendance in the CSG group at a follow-up point of one year."

Reporting Results for Hypothesis 8c:

For both groups compared to each other (AGI vs. CSG)

There is no significant difference between the groups at any time point. The mean scores for AGI group start higher (AGI=93.99, CSG=92.32) and end lower (AGI=87.72, CSG=90.72), but the difference is not significant. Effect sizes favour the AGI with small size effects of 0.18 compared to tiny CSG effects of 0.00 over T2 – T3, and small size effects of 0.31 compared to small CSG effects of 0.13 over T3 – T4. It has to be remembered that these are negative changes taking place with decreasing levels of attendance reported and the combined effect sizes show a greater negative effect happening in the AGI group with a small effect of 0.18 T2 – T3 and a small effect of 0.18 T3 – T4. this difference is not significant.

Summary:

When both groups are considered together there is a significant decrease in the attendance over the T3-T4 follow-up period, an interesting finding, but the lack of any significant difference between the AGI and CSG groups at a follow-up point of one year mean that the experimental hypothesis 8c is rejected and the null hypothesis accepted:

Null Hypothesis 8c:

“According to school attendance records, there will be no measurable difference between the percentage attendance in the AGI group and the percentage attendance in the CSG group at a follow-up point of one year.”

Interpretation of data and discussion of Hypothesis 8c:

In a situation similar to that in Hypothesis 8b, the acceptance of the null hypothesis is inevitable, but the significant findings contained in the data inform the hypothesis and require an account of the effects of group work as presented in this study are to be fully understood. The starting point is the unequivocal finding that there is no apparent effect of the content of group work sessions on school attendance. The significant decline on attendance over the follow-up period raises the question as to whether there is a negative effect of group work per se on attendance considered 1-year post intervention.

In one vital respect an absolute conclusion on this point is impossible with the current data-set, there being no non-intervention control data and no valid matched pairing within the school population to determine naturally occurring attendance patterns over the same period. Certainly to compare the intervention groups with the normal school population is misleading because of the at-risk nature of the children involved, and this is a frustrating factor in analysing the attendance. There is an accepted difference between the levels of attendance in Primary school and those experienced in secondary schools, as discussed above in Hypothesis 6. What is of interest here is that the levels of attendance in the intervention group do seem to be on the decline when measured from T1 to T3 to include the waiting list period (see discussion of Hypothesis 5c).

It is unclear how this effect is echoed in other at-risk populations, but there seems to be a decline over time that is accentuated over the follow-up period during which a high proportion of the intervention group has had a change in school. Hypothesis 6c, in

contrast to Hypothesis 5c and Hypothesis 8c reports a constant level over the Intervention period followed by a sharp decline in follow-up, but the weight of evidence is for a decline over the intervention period, and that it does set a precedent for further decline at T4. The key question for the discussion for Hypothesis 8c is why both research groups show similar deterioration in attendance levels, given the different content of the sessions and the different processes planned for the groups. It revolts common sense to say that the benign nature of the interventions produces pervasive negative effects intentionally, but can epidemiological effects on attendance for at-risk samples simply account for the joint deterioration? Interestingly, the trends in the data show the AGI group attendance falling more sharply than the CSG (92.32 – 87.72 compared to 92.03 – 90.72 for the CSG group), suggesting at the very least a vestige of differential treatment effect.

Attrition:

The overall figures of children present at T4 who were present at T3 have fallen to n=79, giving n=40 in the AGI group and n=39 in the CSG group. The figure of n=79 is 60% of the research group and represents a significant level of attrition over the follow-up period. Children moving schools, exclusions and moving home, can account for the rate of attrition.

Section 5.9: Hypothesis 9

Question 9. Does the attitude of the school to the Intervention and research process affect the outcomes?

If you allocate a score to the schools involved based on a number of criteria relating to the understanding of and interest in the intervention and AGI research programme will you notice those with higher scores having most of the best outcomes? Will one intervention condition be more sensitive to these conditions than the other?

The emphasis given to group factors and school factors in section 3.1.1 draws out their importance in establishing group norms for behaviour and the consistency of condition necessary to develop themes of study over a number of weeks. The support of the teaching staff is considered to be vital in contextualising the interventions and a sympathetic policy with regard to the provision of facilities is thought to be vital to the generation of positive behaviour outcomes. For these reasons, the prediction is that a

positive attitude of the school hierarchy, and a facilitating approach to establishing conditions within the school will have a positive effect on behaviour.

Hypothesis 9:

Positive behaviour outcome for both AGI and CSG groups will correlate significantly with the attitude of the school to the Interventions and research process, measured by scoring within-school conditions.

The Hypothesis is split into three parts (a – c) each relating to a different dataset.

- a) When schools are ranked for outcome using a gain score (T2-T4) from child self-report (YSR), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions
- b) When schools are ranked for outcome using a gain score (T2-T4) from teacher report (TRF), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions
- c) When schools are ranked for outcome using a gain score (T2-T4) from parent report (CBC), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions

Table 5.9: Correlation between rank orders of School Commitment and Gain Score (T2-T4)

YSR

	School Commit	Whole Grp gain score	AGI gain score	CSG gain score
	Rank Ord	Rank Ord	Rank Ord	Rank Ord
STC	1	5	2	5
MWP	2	1	1	3
CEN	3	3	4	2
JSP	4	4	5	4
PSJ	5	2	3	1
Correlation		-0.30	0.60	-0.70

TRF

	School Commit	Whole Grp gain score	AGI gain score	CSG gain score
	Rank Ord	Rank Ord	Rank Ord	Rank Ord
STC	1	3	2	3
MWP	2	2	3	2
CEN	3	5	5	5
JSP	4	1	1	1
PSJ	5	4	4	4
Correlation		0.10	0.20	0.10

CBC

	School Commit	Whole Grp gain score	AGI gain score	CSG gain score
	Rank Ord	Rank Ord	Rank Ord	Rank Ord
STC	1	5	5	5
MWP	2	1	1	1
CEN	3	2	2	3
JSP	4	3	4	2
PSJ	5	4	3	4
Correlation		0.00	-0.21	-0.10

Hypothesis 9a

This part of the Hypothesis reports on findings from the Youth Self report

Experimental Hypothesis 9a

“When schools are ranked for outcome using a gain score (T2-T4) from child self-report (YSR), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions

Null Hypothesis 9a

“When schools are ranked for outcome using a gain score (T2-T4) from child self-report (YSR), there will be no significant correlation for AGI and CSG groups with schools ranked for within-school conditions”

Reporting Results for Hypothesis 9a:

There is a slight negative correlation (-0.30) for the groupwork conditions considered together, a stronger positive correlation (0.60) for children receiving the AGI intervention and a negative correlation of the same order (-0.70) for the children receiving the CSG intervention.

Summary:

Although the correlations do register using the Spearman test, none is strong enough to register as significant. The lack of any significant correlation between the rank orders for gain score for intervention AGI, CSG and AGI+CSG and the order of school ranked for conditions means that the experimental hypothesis is rejected and the null hypothesis accepted:

Null Hypothesis 9a

“When schools are ranked for outcome using a gain score (T2-T4) from child self-report (YSR), there will be no significant correlation for AGI and CSG groups with schools ranked for within-school conditions”

Interpretation of data and discussion of Hypothesis 9a:

Although no correlation of any significance is reported by this analysis and the prediction that the school conditions will affect outcome is rejected, the trends, although contradictory are worth commenting on: When the children assess their own improvements, the CSG group has a negative correlation with school conditions, they think that the worse things are in school according to the practitioner the more likely they are to show positive behaviour. The AGI group on the other hand, has a commensurate positive correlation: they see themselves improving when the conditions in school are right according to the practitioner. This could be highlighting the different style of the interventions offered, with the group factors more important for the establishment of trust and therapeutic gain in the AGI group, or more simply that the children do not recognise the same conditions as being positive within school. Although not significant, the strong opposite correlation is striking, pointing to the conclusion that children from schools rated highly for internal conditions who have the AGI Intervention are more likely to respond to positive conditions when rating themselves.

Hypothesis 9b

This part of the hypothesis reports on findings from Teacher Report Form (TRF)

Experimental Hypothesis 9b

“When schools are ranked for outcome using a gain score (T2-T4) from teacher report (TRF), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions”

Null Hypothesis 9b

“When schools are ranked for outcome using a gain score (T2-T4) from teacher report (TRF), there will be no significant correlation for AGI and CSG groups with schools ranked for within-school conditions”

Reporting Results for Hypothesis 9b:

Analysis reveals a negligible correlation (0.10) for the groupwork conditions considered together when scored by teachers. This lack of significance is echoed in the analyses for the conditions considered separately, with the AGI intervention correlating at (0.20) and the CSG at (0.10).

Summary:

The lack of any significant correlation between the rank orders for gain score for intervention AGI, CSG and AGI+CSG and the order of school ranked for conditions means that the experimental hypothesis is rejected and the null hypothesis accepted:

Null Hypothesis 9b

“When schools are ranked for outcome using a gain score (T2-T4) from teacher report (TRF), there will be no significant correlation for AGI and CSG groups with schools ranked for within-school conditions”

Interpretation of data and discussion of Hypothesis 9b:

The teachers do not see any connection between the practitioner criteria of positive school conditions and the positive outcome of children from either group. This is slightly disappointing as the strong associations of the children with the school criteria are not present. This lack of correlation could mean either mean that school condition makes no difference to outcome, or that the criteria have no resonance with teaching staff.

Hypothesis 9c:

This part of the Hypothesis reports on findings from parent report: the Child Behaviour Checklist

Experimental Hypothesis 9c

“When schools are ranked for outcome using a gain score (T2-T4) from parent report (CBC), there will be a significant correlation for AGI and CSG groups with schools ranked for within-school conditions”

Null Hypothesis 9c

When schools are ranked for outcome using a gain score (T2-T4) from parent report (CBC), there will be no significant correlation for AGI and CSG groups with schools ranked for within-school conditions”

Reporting Results for Hypothesis 9c:

Analysis reveals no correlation (0.00) for the groupwork conditions considered together when scored by the parents of the children, a negligible negative correlation (-0.21) for children receiving the AGI intervention, and a negligible positive correlation (0.10) for children receiving the CSG intervention. This lack of significance is echoed in the analyses for the conditions considered separately

Summary:

The lack of any significant correlation between the rank orders for gain score for intervention AGI, CSG and AGI+CSG and the order of school ranked for conditions means that the experimental hypothesis is rejected and the null hypothesis accepted:

Null Hypothesis 9c

“When schools are ranked for outcome using a gain score (T2-T4) from parent report (CBC), there will be no significant correlation for AGI and CSG groups with schools ranked for within-school conditions”

Interpretation of data and discussion of Hypothesis 9c:

The parents do not see any connection between the practitioner criteria of positive school conditions and the positive outcome of children from either group. This is not

entirely surprising considering that the parents are the furthest removed from the school environment and the as such the furthest removed from the interventions. This lack of correlation could mean either mean that school condition makes no difference to outcome, or that the criteria have no resonance with parents.

Chapter Six: Interpretation and Discussion of Findings

Introduction and Summary of Findings

The preceding chapters have spent some time setting out the key features of the AGI study, its roots and resonances in the literature and the conditions under which data were gathered and analysed. As the findings are gathered and the impact of the interventions assessed in the following chapters, it is timely to be reminded that this study has been directed to the relief of children suffering the self-destructive effects of disaffection. Disaffection is simply a word, but its effects are felt in situations that were outlined at the beginning of Chapter One: the classteacher with the fighting children on a hot day, the junior doctor with the ADHD child, the sad loners and truants sinking in a world of drugs. The urgency of the need, and the sympathy we might feel for the children or their carers must not distract us from dispassionate evaluation of the how the interventions offered in this study have worked. At the same time these are not abstract principles being investigated, but complex social interactions at the centre of which are young people with behaviour strategies and emotional needs that leave them confused and damaged as they negotiate their social environments. As this piece of work nears its conclusion, the image of the child at the centre of its personal maelstrom should be our central reference, not to excuse negative findings (there are some of those), but to remind us that in an imperfect world, imperfect interventions are often the best we can offer.

Even though imperfect interventions are expected to lead to imperfect predictions, nevertheless there is a sense of frustration in the mind of every researcher when his or her predictions are not supported by the data. This is certainly the case in the AGI study. There can be a tendency in such situations to simply ascribe positive findings to the interventions and negative findings to extraneous factors. In the sections that follow, there is indeed an attempt to find explanations for the findings, and these explanations will be more detailed in cases where predictions are not supported by the data. It is hoped that this approach will not be mistaken for unacknowledged researcher bias (issues of bias and reflexivity are

addressed in section 6.4), but is more a reflection of the implications on the literature from which the predictions are drawn. If the predictions formulated into AGI hypotheses have genuinely been drawn from the literature, then unsupported hypotheses will need to reflect back to the literature, and if this is to happen they will need to be examined for extraneous factors. The same process will indeed be followed for positive findings, and extraneous factors promoting false positive findings have an account, for example, in the examination of regression to the mean analyses below, as well as within the individual interpretations. For reasons discussed above and given an account in section 6.4, extra effort will be spent seeking explanations for negative findings.

In Chapter Two predictions were made as to the nature of change expected from Interventions, and the findings from the AGI study have fulfilled the predicted direction of change in twelve out of twenty hypotheses using the questionnaire data and in five out of fifteen hypotheses using the observation and attendance data. This far the interpretations offered have focussed on individual sub-hypotheses and been speculative in nature – looking for explanations where predictions have been frustrated. In this section the speculation is weighed up against the data with the aim of providing cool but fair assessment of the messages to take home from the results. Whether the data supported the hypotheses in the direction of their predictions are summarised in the tables below.

Table 5.1a: Null or Experimental Hypothesis Acceptance 1-4 & 9

	Hypothesis 1	Hypothesis 1(sub)	Hypothesis 2	Hypothesis 3	Hypothesis 4	Hypothesis 9
YSR	Exp	Null	Exp	Null	Null	Null
MSCS	Exp	Null	Exp	Null	Null	-
TRF	Exp	Exp	Exp	Exp	Exp	Null
CBC	Exp	Exp	Null	Exp	Null	Null

Table 5.1b: Null or Experimental Hypothesis Acceptance 5-8

	Hypothesis 5	Hypothesis 5(sub)	Hypothesis 6	Hypothesis 7	Hypothesis 8
On-Task	Exp	Exp	Null	Null	Null
Off-Task	Null	Null	Null	Null	Null
Attendance	Exp	Exp	Null	Null	Null

Table 5.1c: Effect Sizes for Hypotheses 1-4

		Hyp1.	Hyp1(sub).		Hyp2.	Hyp3.		Hyp4.	
		T1-T3	T1-T2	T2-T3	T2-T4	AGI T2-T3	CG T2-T3	AGI T2-T4	CG T2-T4
YSR	TTS	0.82	0.52	0.22	0.44	0.28	0.34	0.42	0.46
	INTS	0.98	0.56	0.38	0.55	0.35	0.42	0.5	0.61
	EXTS	0.39	0.34	0.02	0.27	0.14	0.2	0.2	0.34
MSC S	TTS	0.66	0.59	0.06	0.61	0.25	0.34	0.48	0.76
	AcTS	0.4	0.41	-0.02	0.39	0.16	0.31	0.49	0.28
	AfTS	0.57	0.56	-0.03	0.6	0.37	0.26	0.51	0.7
	CoTS	0.71	0.52	0.19	0.57	0.41	0.23	0.57	0.56
	FaTS	0.42	0.51	-0.09	0.39	-0.1	0.4	-0.04	0.9
	PhTS	0.59	0.31	0.29	0.23	0.14	0.35	0.06	0.41
	SoTS	0.74	0.58	0.08	0.57	0.36	0.16	0.68	0.45
TRF	TTS	0.58	0.05	0.5	0.64	0.73	0.15	0.84	0.49
	INTS	0.55	0.09	0.45	0.73	0.53	0.05	0.92	0.57
	EXTS	0.39	-0.01	0.41	0.45	0.61	0.25	0.57	0.34
CBC	TTS	0.28	-0.03	0.31	0.04	0.43	0.26	-0.07	0.13
	INTS	0.15	0.61	0.15	0.21	0.48	0.06	0.13	0.28
	EXTS	0.28	-0.04	0.34	-0.11	0.25	0.23	-0.15	-0.07

Table 5.1d: Effect Sizes for Hypotheses 5-8

	Hyp1.	Hyp1(sub).		Hyp2.	Hyp3.		Hyp4.	
	T1-T3	T1-T2	T2-T3	T2-T4	AGI T2-T3	CG T2-T3	AGI T2-T4	CG T2-T4
C-ob: a	1.06	0.28	0.77	0.00	0.15	0.15	0.38	0.06
C-ob: b	0.12	0.12	0.24	0.51	0.01	0.08	0.59	0.42
Attend	0.19	0.16	0.31	0.33	0.21	0.28	0.46	0.14

Table 5.1e: Attrition

	Hyp 1: n= T1 – T3	Hy1(sub) n= T1 – T2	Hyp 2: n= T2– T3 – T4	Hyp 3: n= T2 – T3		Hyp 4: n= T2 - T3 – T4	
YSR	44	44	77	AGI	45	AGI	39
				CSG	43	CSG	38
MSCS	28	28	75	AGI	42	AGI	37
				CSG	41	CSG	38
TRF	82	82	104	AGI	58	AGI	51
				CSG	59	CSG	53
CBC	31	31	60	AGI	38	AGI	28
				CSG	39	CSG	32
	Hyp 5: n= T1 – T3	Hy5(sub) n= T1 – T2	Hyp 6: n= T2– T3 – T4	Hyp 7: n= T2 – T3		Hyp 8: n= T2 - T3 – T4	
On-Task	16	16	43	AGI	45	AGI	21
				CSG	43	CSG	22
Off-Task	16	16	31	AGI	36	AGI	16
				CSG	35	CSG	15
Attendance	84	84	79	AGI	52	AGI	40
				CSG	51	CSG	39

Section 6.1: Interpretation of Hypotheses

6.1.1: Hypothesis 1

This hypothesis is framed in two parts: the first predicting Improvement over the entire period (T1 – T3), the second (labelled 'sub') predicting that the change over the intervention period (T2 – T3) to be greater than the waiting-list period (T1 – T2). In all of the divisions a-d of the hypothesis, significant change was found over the entire period, but an interesting diversion of scores happens when considering in which time period (T1-T2 or T2-T3) this occurs. The teachers report convincingly, with very high significance (0.000 in Total, Internalising and Externalising scores) that change happens in the Intervention period that is preceded by a period of no-change. The parents agree with the teachers, but self- reported change is less conclusive. Parents do not report any significant change over the waiting-list period, going on to report significant Externalising gains in the intervention period, no significance in Internalising scores and very close to significant change in total scores ($p=0.056$). Both teacher and parent report are at odds with the two self-report measures, which report significant gains over the waiting list period followed by no reported change over the intervention period. The children's reports are quite clear about this: change in the waiting-list period is highly significant, change in the intervention period is not close to being significant. The exceptions to this trend are the Internalising T-scores on the YSR which as well as being significant T1-T2

are also significant T2-T3, while the other exception is the physical sub-scale on the MSCS which is significant neither T1-T2 nor T2-T3. The presence of these anomalies at the very least shows that the effects are not general in the sense of there being an unthinking response from the children, and at best throw light on the divergence in the reports. Before advancing explanations for the existence of gains in the waiting list period and the lack of gains in the intervention period in the self-report measures, it is worth considering the nature of those gains:

The self-report data give clear statements about the nature of change over the period T1-T2: the scores improve and this gain is significant. The data also give clear statements about the nature of change over T2-T3: the scores marginally improve but the gains are not significant. There are also data relating to the magnitude of change over these time periods, the effect size measuring the change in terms of the standard deviation of the sample population. In all the self-report data the effect size is larger for the waiting list period T1-T2, but is it significantly larger? Because of the statistical impossibility of comparing two effect sizes, a supplementary analysis was conducted to discover whether the significant gain of T1-T2 was of significantly greater magnitude than the non-significant gain of the T2-T3 intervention period. This analysis was conducted to test a post-hoc hypothesis that looks for evidence that the data can be brought in line with the predictions made in Hypotheses 1a, 1a(sub), 2a and 2a (sub). **This post hoc hypothesis states that the magnitude of gain in the intervention period will not be significantly lower than the magnitude of gain in the waiting list period.** In other words, there will be no difference between the amount of change found in either of the time periods. The gains/losses during the waiting list period were compared with those during the intervention period by computing the gain scores for each variable:

Gain 1= T1 mean score – T2 mean score (waiting list period)

Gain 2= T2 mean score – T3 mean score (intervention period)

Paired t-tests were performed between Gain 1 and Gain 2.

Table 6.1.1: Post-Hoc Waiting List Analysis

Paired samples statistics				
	N	mean	Std. deviation	Significance
Gain 1	44	5.57	12.05	-
Gain 2	44	2.45	11.36	-
Paired samples t-test				
Gain 1- Gain 2	44	3.11	20.73	0.325

These results provide evidence to support the post hoc hypothesis. It seems that there is no statistically significant difference between gains made in the waiting list period and those made in the Intervention period. This brings the self-report data more into line with the predictions made in hypothesis 1, although it does not fully explain the divergence found in the data when teacher and parent reports are considered.

Explanations for this divergence can be found in a number of sources:

The effect of selection, expectation and anticipation relating to a positive intervention.

The effect of wanting to present the best possible aspects of self: 'presentation' or 'conformance' effects.

The effect of generalised population gains.

The effect of administering a questionnaire leading to self-questioning and reflection.

The explanations and interpretations of the data have focussed on a 'selection effect' alongside a 'conformance effect' to perceived expectations on entry requirements into the groups. The data do not allow informed speculation on the generalised population trends. It is possible that there are gains in Internalising and Externalising symptoms that are the result of natural maturation, but as the sample is not homogenous in age, and were assessed at different points in the year, it seems unlikely that this effect can account for the data. The effect of administering a questionnaire as an explanation of the data trends is likewise discounted because of the gains are not found throughout the second administration.

The selection effect is explained thus: the notion of 'Health' might be thought to include the regular, if not necessarily frequent, feeling of wellbeing e.g. Maslow (1968). The feeling of wellbeing is thought to include an oceanic feeling of being a functioning part of a greater whole, a feeling that the feeler's contribution however small is not insignificant.

Selecting children screened to include those susceptible to feelings of confusion and hopelessness and informing them of their inclusion in a benign intervention extraordinary to their experience could have the effect of replicating the sense of purpose underlying feelings of well being, at least in the short-term. It is interesting that these putative improvements in self-image and self-reported problems are sustained over the intervention period, which suggests that expectations must have in some way been met. It is interesting also that the Internalising problem scores (YSR) continued to improve T2-T3, suggesting that there is an internal process at work in their behaviour change. The change in perception and behaviour does not translate into how they rate their physical selves, and this fact provides a reassuring reference to the interior nature of this change.

The conformance effect is more straightforward and is proposed to consist of the children, eager to be included in the interventions, scoring their sheets in the way they think is most likely to lead to their inclusion. In this case the scores at T1 are taken to be a more or less accurate representation of their problems and self-image, there being little time to think about the consequences. The second assessment, done 12 weeks after the first, during which time the children will have been experiencing this 'selection' effect while experiencing anxiety wondering when the intervention will occur, provides the children with the opportunity to score the papers, which they now have familiarity with, in the way that they think makes them less problematic and more likely to be included in the groups.

The net effect of these two proposed processes would be to lead children to report change in the waiting-list period. The processes are essentially interior and cannot reasonably be expected to lead to behaviour change, indeed they may result in confused behaviour as the expectation of the selection effect is frustrated over time.

6.1.2: Hypothesis 2

These analyses give the overall picture of improvement during intervention (T2-T3). Teachers (TRF) report a continued improvement on Internalising and total problem scales during follow-up (T3-T4), children (YSR) report no significant decline and parents report a significant deterioration in Externalising and problem scores. At the same time

children rate their self-esteem (MSCS) continuing to improve with gains made on total scores, affect, competence and social sub-scales. According to questionnaire data, children are displaying fewer emotional problems in school as time goes on while behaviour problems sustain improvements. This view of increasing emotional competence is supported by the gains made over the same period in children's perceived self-ability to manage emotional situations. Over the same period parents see children's behaviour deteriorating after initial improvements and no significant impact over any period on emotional problems.

It is possible that children, keeping the lid on behaviours at school while undergoing a continuing process resulting in a reduction in emotional problems, find an outlet at home. This supposition is plausible and leads to what might be called a "displacement effect" whereby improvements at school result in raised expectations from teachers and positive peer pressure both of which support behaviour gains but neither of which is present at home. The damaging behaviours that no longer have a place at school, re-surface at home.

In addition, it is perhaps telling that the parents do not see Internalising problems improving over any period. This might reflect a relative lack of concern for Internalising issues that generally result in quiet behaviours, or parents' own lack of emotional aptitude (the same could be said for teachers, but they have been more closely involved with the children's educative experience in groupwork). In other words, it is reasonable to expect parents to notice behavioural problems or to attribute emotional symptoms to behavioural problems.

The other trend worthy of attention is the comparison between the self-reports of Hypothesis 2 with those of Hypothesis 1 in as far as they both report data that overlaps in period T2-T3. In discussion of Hypothesis 1 no significant change was reported, while hypothesis 2 reported significant change sustained at follow-up T4. There is an apparent incompatibility in the findings from these two hypotheses that warrants some attention: 1a and 1b proposed the existence of a "selection" effect allied to positive mental health to explain gains not attributable to intervention – and the absence of those gains from the intervention period. Hypotheses 2a and 2b found gains in the intervention period and made some attempt at reconciling these findings with those in 1a and 1b. In essence,

this reconciliation focussed on the relatively low numbers of self-reports relating to the pre-intervention period (T1-T2) and proposed that effectively two different populations had been constituted: Population 1 present at T1, T2 & T3 (n=27), Population two present at T2, T3 & T4 (n=77). As well as the increased statistical power in the increased size of Population 2 which would boost the non-significant gains made over the intervention period T2 – T3 in Population 1, a large proportion of Population 2 were not assessed at T1, for these children their “selection” effect would operate during the intervention period – enhancing gains made from treatment.

In addition to the statistical significance found in Hypothesis 2a the effect sizes are larger than those found in Hypothesis 1a: YSR internal and external T-scores are 0.55 and 0.27 for Hypothesis 2a while they are 0.38 and 0.02 for the same period (T2-T3) in Hypothesis 1a. Effect size calculations measure magnitude of change independent of sample size and this seems to suggest that children do better if they have not received the questionnaire in the waiting-list assessment. If this were true it would support the proposal of effects of selection and conformance in as far as they could be seen to be active from T2 for the majority of children.

6.1.3: Hypothesis 3

This hypothesis looked for differential gains between the AGI and CSG groups over the intervention period T2-T3. The evidence was weighed in Chapter 4, and it was felt that the effect size analysis gave a better basis for comparison, because effect sizes operate on the same scale, independent of sample size. Differing values between effect sizes for the intervention groups can be reported when substantial difference is found. P-value analysis is not a uniform scale as it varies with sample size, and is used as a supporting analysis when the sample size is low (for reasons of attrition and splitting the research population into two samples – AGI and CSG).

With this in mind, the data from the children do not distinguish between the groups, but both parents and teachers report an effect of intervention favouring the AGI group. The teacher data shows the magnitude of change categorically favouring the AGI group on both the Internalising and Externalising problems sub-scales: this is supported by statistically significant difference in the ANOVA test. The parent data shows the

magnitude of change categorically favouring the AGI group on the Internalising problem sub-scale, and a parity of effect size in the Externalising scale. This is not supported by statistically significant difference in the ANOVA test, but weighing the magnitude of change, it was felt that this difference between the groups, uniform across contexts was worth reporting as positive.

These findings provide encouragement for the proposition that the content of the group sessions is a decisive factor in the nature and extent of gains made. However it would be premature to leap to this conclusion. Both child reports; the problem questionnaire and the self-concept report no difference between the groups, and while there is no need to repeat the interpretation made in Chapter 3, an explanation of the way this hypothesis is split between null and experimental hypotheses is offered below.

An explanation could lie in consideration of the content of the two group interventions: both seem to have achieved the same result (at least in the short-term) but the approach has been different. Common elements will be the atmosphere of adult concern and attention, the encouragement of mutual respect, positive feedback and containment of behaviours. The AGI group goes further to make explicit reference to the emotional base of behaviours, seeking to instil a reflective ability, and an ability to emotionally self-regulate. It seems that what are being reported in Hypothesis 3 are the effects of the benign humanistic approach and positive group process that underpin both interventions. In that the children report the effects of positive feedback and adult attention with effects on self-image and reduction in problem behaviours. The fact that parents and teachers notice these improvements more in the AGI group gives some credence to the critique of self-concept offered in Chapter 3: there can be a gap between perception and reality.

Another possibility is the notion of a so-called 'sleeping effect' noticed by Kolvin when he reported change emerging after a number of months. It could be that the way this hypothesis is split is a product of the nature and process of behavioural change, and it will be resolved as data from the follow-up point are discussed.

6.1.4: Hypothesis 4

Analysis of the data show gains made over the intervention period T2-T3 sustained at a follow-up point of 1 year (T4) when both AGI and CSG groups are considered together. This is an important finding, detailed in Hypothesis 2 and one which advocates strongly for the use of groupwork in the school setting as an aid to socialisation, and a modifying factor for problem behaviours.

Hypothesis 4, however, is aimed at distinguishing between the intervention conditions of Action GroupSkills and Curriculum Studies. The data are examined on the same basis as Hypothesis 3: looking first at effect size and then at ANOVA. The major positive finding is that teachers (for the first time blind to treatment status), report differential gains over the period T2-T4 favouring the AGI group. The magnitude of change in the AGI group is reported to be approximately double that of the CSG group on both Internalising and Externalising sub-scales. This is not supported by statistical significance, but for reasons given above, emphasis is given to the effect size analysis and this is a major finding from this study, not often found in the literature.

Data from children and their parents do not distinguish between the groups at T4, one-year post intervention. The conclusion on these data with this set of experimental methodologies points to an assertion that the AGI intervention is more effective in reducing problematic behaviours and promoting self esteem than CSG groups focused on curriculum matters over the long-term. But this differential effect is only present in school.

Looking for explanations as to why the hypothesis is split between respondents in this way, one might look to methodological concerns addressed elsewhere (namely: sample size, the matching of groups, procedural controls in administration of questionnaires and so on), and to more to sociological explanations as to the context-specific nature of behaviour.

There are two minor anomalies, the Family Scale on the MSCS and the Internalising scale on the CBC, both of which distinguish the CSG as making gains over the AGI group in a way counter to the predicted effects. From these anomalous findings it is possible to discern a trend for AGI children to appear more emotionally problematic to

their parents, and CSG children to rate familial relations more positively. The parent data itself is anomalous in that parents are the only respondents who report a tendency toward deterioration over the follow-up period. The total scores show a symmetrical pattern of moderate gains T2-T3 followed by losses of comparable magnitude. The Internalising scale shows significant difference between the groups, and reason for this is clear from the data, the parents report a 3-point gain over intervention for the AGI group followed by losses that leave the children slightly improved. The CSG group makes no gains over intervention, but parents report a three-point gain over the follow-up year. The children themselves reporting in the MSCS rate their improvement overall as shallow, gradual and equal for both groups, the family sub-score shows a slight loss for the AGI group T2-T3 (47.92-48.97) that remains unchanged at T4. At the same time the CSG group makes a slight improvement at T2-T3 and another T3-T4 (52.92 - 49.82 - 46.13). The opposing direction of these slight changes produces the difference resulting in significance. These isolated findings, restricted to sub-scales, are by themselves little more than a snag in the data for which it is difficult to find satisfactory explanations. If however they were to be substantiated in any way by analyses for later time points they would present deeply perplexing problems for interpretation.

6.1.5 Hypothesis 5

Hypothesis 5a looked at the time spent on-task in the classroom situation, and it is here that the data confirmed the predictions made: there was no change over the pre-intervention period (T1-T2), followed by a significant increase in time observed on-task over the intervention period (T2-T3). Hypothesis 5b looked at observed off-task behaviours compared to the rest of the class, and revealed no significant change over any time period. Hypothesis 5c looked at attendance and found no significant change over the pre-intervention period, but also found children significantly less likely to attend school over the intervention period (T2-T3), a finding in the opposite direction to the change predicted. This presents a mixed bag for analysis, and while it is possible to emphasise the findings in keeping with the research predictions, there is little to be gained in the search for truth by doing this at the expense of the analyses not in line with the research predictions. The off-task ratios show no impact of intervention, and may be explained by methodological inadequacies in that it is possible that the behaviours that the observation is seeking to measure are not the behaviours most susceptible to

change, at least during the intervention period. Off task behaviours are likely to be entrenched responses to situations beyond the control of the intervention: one might predict that a common feature of these responses could be a frustration with the delivery of the curriculum, with behaviours masking a lack of competence or interest. Dealing with these sorts of behaviours over the short period the data covers is a tough call. A significant factor in the discussion of these two sets of observation data is the very low numbers of children for whom observations were completed. This would mitigate against finding significant change (as is the case for Hyp.5b) or it could throw up significance (as is the case in 5a) unrepresentative of the larger research population. One has to therefore treat the acceptance of the experimental Hypothesis 5a with some caution and not be surprised by the lack of significance of Hypothesis 5b. No explanations of low numbers can be used to account for the really perplexing finding from this Hypothesis set: that of children bunking off school during the intervention period. Hypothesis 5c looked at attendance and found that while there was no significant change during the pre-intervention period, there was highly significant deterioration of attendance while the interventions were taking place (T2-T3). With attendance of groupwork sessions at 98%, this would mean other lessons on other days suffering. This is counter to the research predictions which look for a positive effect on school engagement as a result of the increased attention of the small group and the increased confidence that this is predicted to engender. The finding cannot be convincingly explained and we look for the data to be illuminated in the discussion of subsequent hypotheses.

6.1.6 Hypothesis 6

The general hypothesis that observable gains in the intervention groups (AGI+CSG) are sustained at one year post intervention was supported by the evidence from one of the three data-sets: the observed on-task behaviours showed no significant change over the follow-up period T3-T4. The evidence from the other two datasets did not support the experimental hypotheses and for both the observed off-task behaviour data and the attendance data, the null hypothesis was accepted. Interpretation of these findings leads to the assertion that the intervention group spends a high percentage of time on task and this classroom application is present one year after the intervention has finished. The on-task behaviours for this group showed significant gains over the intervention period T2 – T3 compared to the waiting list period T1-T2 and these gains are indeed sustained at

follow-up. There seems to be a real effect of increased time spent concentrating on their work and this is to be celebrated.

This positive finding does not translate into positive effects on non-aggressive off-task behaviours and attendance, the first of which might certainly be expected under the circumstances. These non-aggressive behaviours might be characterised as those incidents of interaction between children that are not related to the task offered by the teacher. They are not necessarily problem behaviours in themselves, but are likely to be thought of as such by the teacher in the task of completing a lesson plan. The inability to resist distractions therefore, does not necessarily negate the finding that this population spent significantly more time on task at one year than they did prior to intervention. The non-aggressive behaviours are measured in relation to the rest of the class, so do not indicate an objective assessment of frequency. It might be argued therefore that non-aggressive classroom behaviours are a way of compensating for increased time spent on-task, that there is a natural correspondence, particularly for restless boys between increased time on task and a perceived increase in non aggressive behaviours relative to the rest of the class as the children 'let off steam' after periods of more intense concentration. For Internalising children not naturally disposed to making a fuss, the increased non-aggressive behaviours may be a sign of their increased social confidence, and ability to actively manage the demands of schoolwork with the benefits of being a mover and shaker in the classroom, a role that will demand more visibility and therefore more observation.

It seems indisputable that children involved in school and engaged in their education, able to engage positively with their peers and their teachers (all aims of the interventions offered), must be more able and more likely to attend school. The intervention has either not achieved the aims in relation to attendance, or produced positive attitudes that result in non-attendance. If the attitudes produced by the intervention were to be inferred from the questionnaire data gathered in the study, they would be likely to be of a reflective type. Attitudes encouraged by the intervention would predispose the participants to questioning the various inputs they receive before making a considered judgement as to the course of action. If the participants of the groupwork produced young people more independent in their thinking and attitudes, then it could be that a proportion of them are choosing to exercise that independence by making an active choice not to attend school.

But it could equally well be pointed out that attendance is a key indicator of disaffection, and that the phenomenon of the engaged non-attender is a rare one. There is however a strong case to be made for a third explanation, and this rests on the changing attendance patterns from primary to secondary school. The figure of 93% attendance at T2 falling to 89 % at T4 is below that expected for the whole school population from the schools in question (the level is a mean of 94% in secondary and 98% at Primary), but this is an at-risk population to start with and are therefore most likely to be affected by the increased personal responsibility in attending secondary school. It is disappointing that the intervention does not appear to have a more positive effect on the key indicator of attendance, but one wonders if the level would have been even lower had the intervention not taken place. Unfortunately the identification and tracking of a non-intervention control is outside the scope of this study.

6.1.7: Hypothesis 7

There is no finding of significant difference between the groups over the intervention period T2-T3 on either of the observations or the attendance findings. The intention of Hypothesis 7 is to establish if the content of the group sessions produces observable differences in school behaviours, the immediate indicators of disaffection. Contrary to what is predicted, there appears to be no difference between the AGI intervention and the CSG intervention. The question then is why do the findings refute the prediction? While the first instinct is to blame the reliability and sensitivity of the instrument, a more reflective response might look to the nature of change sought and the time frame over which change is sought. The time frame is short and any benefit of intervention would have to have an almost immediate appearance in 'real-world' behaviours. The benefits of groupwork are likely to be characterised as being generated through a feeling of value transmitted from the facilitator to the children and eventually between the children. In addition the AGI focussed on developmental encouragement and reflective abilities related to enacted situations, while the CSG focussed on encouragement of learning outcomes. The more profound change is anticipated from the AGI children, and a lot more is asked of them emotionally during the course of the groupwork sessions. In contrast the CSG group are engaged in activities the like of which they are familiar with, although the warmth and rewards available for performance are unusual if not impossible in the classroom situation. Both these processes might reasonably be

expected to produce beneficial effects in real-world outcomes of classroom behaviour and attendance: the AGI because of increased reflective and relationship skills, the CSG because increased curriculum skills is argued to lead to better integration with school. However, because of the short time span, and because the behaviours targeted in the intervention sessions (reflection etc. for the AGI and curriculum skills in the CSG) could reasonably be expected to require time to register in the outcomes measured in Hypothesis 7, it is perhaps unsurprising that the predicted improvement does not show up.

The reflective response to the findings is that the deep changes in patterns of behaviour sought from the AGI intervention could not be expected to be active over the T2-T3 time period. The absence of change in observable behaviours in the CSG group echoes the reasoning behind the lack of change in the AGI group. The cognitive competence upon which a lot of school performance depends requires patience and change can be anticipated over a long period. It is therefore possible to assert that the impact on behaviour measured over a short time period could only relate to primary processes of transferred value, because the secondary processes activated in both intervention groups require a longer time to register in observable behaviours.

6.1.8: Hypothesis 8

This hypothesis is looking for difference between the groups on observed behaviours and attendance over the follow-up period (T3-T4). The very slight decline in observed on-task behaviours over the follow-up period is not surprising, since the majority of children will have experienced a change of school between T3 and T4: either from primary to secondary or middle to high. The different expectations and culture of the high school can go a long way to explaining the deterioration of concentration in class time, since the emphasis in high school switches to the learner, and self-motivation has a larger part to play in this situation. This explanation however cannot hide the frustrated prediction of improvements in the AGI group exceeding those in the CSG group, the focused developmental groupwork being predicted to deliver benefits to engagement with schooling not matched by small curriculum studies groupwork. The fact that the evidence here shows no such difference might be attributed to the effect of small-group

interventions, regardless of content, a supposition gained from the fact that both groups experience merely a non-significant falling off of on-task percentages T3-T4.

The maintenance of levels of observed on-task behaviours over the follow-up period has to be seen in the light of significant increases in non-aggressive off-task behaviours for both groups in comparison to the rest of the class. In this respect this hypothesis has little to add to the discussion of Hypothesis 6 except to propose differential learning processes that might lead the groups to the position where they were indistinguishable in their observed increase in off-task behaviours but able to maintain on-task levels. These explanations centre around the predominantly cognitive processes of the CSG group not being geared to impact on off-task behaviours. The AGI by contrast has a focus on reflective action and confidence building, the net effect of which is to increase off-task behaviours, while maintaining on-task behaviours. The reasons for doing so, however, are very different to those explanations of the CSG behaviours. Because the focus of Hypothesis 8 is the difference between the groups there is no wish to repeat the discussion of Hypothesis 6, which sought to explain the whole group (AGI+CSG) effects of intervention. It is however worth re-iterating the feeling that the experience of significant increase in observed off-task behaviours related to the rest of the class in both groups could be due to methodological factors:

- a product of having selected problematic children for inclusion in the study
- an observation effect highlighting the behaviours of participating children (AGI+CSG) where the intervention group was being compared to the rest of the class and observer knowledge of participation status could not be avoided.
- no established validity and reliability levels for the Observation (b) protocol

These arguments received an airing in Hypothesis 6 but the fact that there was no difference between the groups cannot be explained by methodological concerns. While acknowledging the limitations of these protocols there is no possibility of using methodology as a get out of jail free card. The data need interpretation, the attendance data too follow the same general pattern as described in relation to Hypothesis 8b, a period of relative stability T2-T3 followed by a significant decline in attendance. Again the role of changing school is discussed in Hypothesis 6, but again in frustration of predicted direction of change, there is no difference between the AGI and CSG groups. The Hypothesis 8 made educated predictions based on an understanding of the literature, and the lack of evidence for difference between the groups in the data must

lead either to questions about the power of the experimental design to detect change, or an incomplete appreciation of the role of context (in this case school), as an agent of change. In this latter case, a fuller reading of the literature might have resulted in interventions being geared to assisting the classteachers to embed behavioural change that resulted from input outside the classroom.

6.1.9: Hypothesis 9

This hypothesis looked for a correlation between positive behavioural outcome and conducive within-school conditions. This prediction was made because the group and school conditions were considered so important to the running of the groups. No significant correlation was found using the TRF, YSR or CBC. In addition, when the whole group was split between those receiving the AGI intervention and those receiving the CSG intervention, no significant correlation was found. The null hypothesis was accepted in every case. These findings that are counter to expectation and prediction can be interpreted according to two main considerations: the prediction was wrong and the school conditions play little part in determining the effectiveness of the group or the school conditions were inaccurately measured.

Considering first that the prediction was wrong and the school conditions play little part in determining the effectiveness of the group. The commonsense case for spending time establishing baseline conditions for work to take place is strong, and the how-to literature invariably reserves a place for practical considerations. Out of a sense of courtesy to participants and stakeholders, spending time on these matters is important, and that is before the importance of contextualisation of intervention gains in the environment is considered. If these considerations were to be overturned it would require some persuasive evidence. Looking at the gain scores for teachers and parents there seems little room for alternative interpretation, both see little or no relationship between the school environment and improvements in children's behaviours. For this to happen, the teachers and parents from schools scoring poorly on the within school measures criteria will rate the improvements more or less equally to those teachers in schools whose rating was higher. AGI children report their gains to be in some way related to positive school conditions and CSG children report gains to be related to low scores on school criteria. Too much interpretation of sub-significant trends is a danger here and it could be that the interventions were of such a strength that they overcame the limitations of

circumstance, an attractive argument, advocating the power of encounter over the situation in which the encounter takes place.

We need now to consider whether the school conditions were accurately measured. The data used to construct the rank order of the schools were non-parametric and not normed across a population. There are legitimate questions to be asked about the reliability and validity of scores allocated on the basis of a single practitioner's impressions of school process. Nevertheless there is a counter-argument to say that few others were able to make the kind of judgements required and that some efforts were made to address the inherent bias, evidence for which is the lack of support for the experimental hypotheses. On balance it can be considered to be a flawed, but useful measure of school condition and the notion that it is entirely useless as a measure can be rejected. If this is the case and the rank order presented represents with a measure of accuracy the degree to which the schools supported the research interventions through internal and external processes, then there is a question about the interval of the scores allocated. Each school was subject to an involved research process of communication and persuasion at all levels of the hierarchy, the outcome of which was conditions judged to be suitable for the carrying through of the research interventions. It is entirely possible that all the schools subscribed to a minimum level of supporting conditions beyond which the measures used are not sensitive enough to distinguish. In other words, all schools developed adequate processes to support the interventions. Beyond that the criteria for successful behaviour outcome on a mixture of intervention factors such as engagement of subjects together with 'advanced' school processes, but the measures used are not sensitive enough to distinguish between the two.

It seems reasonable to assert that a combination of flawed interval data and interventions independent of conditions above a certain baseline can account for the acceptance of the null hypothesis in the predictions of Hypothesis 9. This leaves intact the strongly attested importance of baseline factors while asserting the strength of intervention.

Section 6.2: Explanation and Synthesis

6.2.1: Explanation

The overall pattern of the research findings is both encouraging and frustrating. The research team has predicted change due to intervention measured by questionnaire and non-questionnaire data. The questionnaire data, established in reliability and validity, provided evidence for 20 sub-hypotheses. Support was found for the experimental hypothesis in 12 instances in that reported change was in the direction predicted. This equates to 60% success in finding support for predicted effects, and is interpreted as an encouragement for the theoretical and practical case for group interventions and the superiority of the AGI condition.

The non-questionnaire data provided evidence for 15 sub-hypotheses, and found support for the experimental hypotheses in 5 instances. This is a frustratingly low percentage of support for the direction of predicted change, and interpretations of the negative as well as the positive findings have been proposed, both in Chapter 4 and in Section 1 of this chapter. Here in section 2 we look at the overall flow of the data and what the patterns of experimental and null hypotheses have to say about the research topic. Can this big-picture view of findings help in reconciling negative findings with positive as synthesis is attempted?

This section looks for convincing reasons why the pattern of evidence has thrown up a percentage of negative findings worthy of note. The pattern of evidence contained in the tables at the beginning of the chapter is examined before synthesis of what appear to be contradictory findings is offered. So as not to repeat the reporting of findings from Chapter 4, nor the interpretations offered in the first section of this chapter, the findings are first summarised by hypothesis (vertically) and then analysed by instrument (horizontally). In doing so new perspectives are sought and new explanations advanced.

Vertical comparison is only meaningful for the questionnaire data because the respondent groups are all reporting in the same mode. Examination by hypothesis does show clear patterns of agreement between the respondent groups: In the first case, it is possible to see some striking agreement where hypotheses have been supported. The consistent evidence from multiple sources shows that the intervention produces a reduction in problem behaviours, an increase in self-esteem and an increase in time spent on-task in the classroom while the group interventions are taking place (T2-T3).

There is clear evidence that improvements in terms of increased self-esteem and reduced problem behaviours can be detected one-year post intervention (Hyp 2 & 6).

Vertical comparison also shows areas where there is disagreement amongst the respondent groups as to whether the hypotheses have been supported. This grey area includes whether it is possible to detect change distinguishing between the intervention groups (AGI or CSG) over the intervention period (Hyp 3 & 7) and at one year (Hyp 8), also there is striking disagreement on the effect of the waiting list period T1-T2. Finally, vertical comparison shows striking agreement where hypotheses have not been supported, which happens in two instances for the non-questionnaire data, in Hyp 7 & 8.

Comparison by hypothesis therefore shows a clear pattern of evidence that supports the effectiveness of groupwork over no intervention, the benefits of which can be felt one-year post intervention. There is some evidence that working in the AGI group is more effective in producing benefit over the period of intervention, but there is no evidence that the intervention groups can be distinguished at one-year post intervention.

Comparison by instrument looks horizontally across the grid to detect any patterns of support given to the hypotheses over time by the respondent group. The teacher report gives the most consistent support to the experimental hypotheses over all time periods and for all hypotheses. It is no surprise that a school-based intervention should have the biggest effect in school, and it is encouraging to see that the effects persist even when the respondents are blind to treatment status as they are at one year. The children perceive themselves to be benefiting consistently from the beginning of intervention to one year (the post-hoc analysis of Hypothesis 1(sub) reveals there to be no difference in the magnitude of gains made T1-T2 and T2-T3), which is encouraging, but the children do not distinguish between which interventions they received. The parent report gives less consistent support to the predictions made in hypotheses, but given the difficulty in demonstrating cross-contextual gains, this is not perhaps surprising, and it makes positive findings more impressive when found.

Still looking horizontally across the grid, it is possible to detect patterns in the non-questionnaire data that were not possible by comparing each non-questionnaire instrument with each other vertically at each time point. The on-task observation gives

the most consistent support to the experimental hypotheses, with benefits detectable at one year but no difference between the groups. Attendance effects are present up to the end of the intervention period, and off-task behaviour ratios give no support to the experimental hypothesis at any point.

The central task at the end of this piece of work is to answer the primary question of whether the intervention has been effective overall in challenging disaffection, reducing the symptoms of disaffection; promoting self-esteem, classroom conduct and attendance. From this position, the secondary question of whether one intervention is more effective than the other in challenging disaffection can be tackled. It seems clear from an initial interpretation of the patterns that emerge from the data that there is a weight of evidence to support groupwork as effective in challenging disaffection up to the follow-up point of one year. There is a weight of evidence to suggest that it is possible to distinguish between the groups over the intervention period, and only teacher-reported evidence to suggest that it is possible to distinguish at one year.

Before explanations for the patterns presented above are advanced, there is a need to be clear that the patterns offered for explanation show the effects of intervention, and not extraneous variables and biases amongst respondent groups. The methodological procedures followed and the use of validated instruments where possible all aim to reduce these biases, but the patterns in the data can assist in the task of highlighting bias. We want to answer the question of whether the intervention was effective, and so we need to be sure that the instruments used were i) reliable in measuring any change and ii) effective in measuring the change targeted by the interventions. The questionnaire instruments have established figures for validity and reliability, and attendance figures are categorical data, for these instruments there can be confidence that change is reliably reported – but is it the change that the interventions are seeking to produce? Looking at the patterns in the findings gives a different perspective to the question to the analysis given to individual hypotheses. Looking at the patterns formed by the findings from individual hypotheses gives additional insight into critical questions about establishing the primary question of the research – is the intervention effective?

This applies to the off-task observation, which is consistent in the lack of support the data gives to the experimental hypotheses at any time point. As pointed out in Chapter 3

and Chapter 4, validity and reliability have not been established for this particular observation protocol, and there is a question as to whether the methodology of data collection was rigorous enough to bring bias to an acceptable level. It is therefore legitimate to ask here whether the instrument itself is reliable when the pattern of findings from this instrument is considered in relation to the others. At the same time, in the interests of reflexivity, it is worth asking whether the researcher would be making the same point if the findings had been positive. Perhaps not, and it is worth reminding ourselves that it could be the case that the simplest explanation is that the interventions do not produce the kind of effects sought after by the hypothesis.

The dataset does have limitations of methodology and size, and it is tempting to ascribe the difficulty in finding differential effects of intervention to these limitations. Again there is a need for caution if the impact of negative findings is to have a proper account here. Differential findings were predicted, and found in school up to one year and in the home and school over intervention. The lack of differentiation by the children indicates the strength of the CSG intervention (or the weakness of the AGI intervention) and was not fully anticipated in the generation of the research questions and the subsequent predictions. It also could indicate the apparent indifference of the children to the content of the group sessions. That the teachers were blind to treatment condition at one year adds status and credibility to the assertion that group content was decisive, but the lack of generalisation of differential gains outside the school situation is something we should be wary of explaining away too easily.

Despite their limitations, these positive findings are not duplicated in many published studies. If they are to have any weight in the support they give to the reading of the literature on child disaffection given in Chapter One, these positive findings need examination in the same way as negative findings invite explanation. The possibility of type 1 errors is pervasive and there is the possibility of random positive findings thrown up by multiple analyses. In order to secure the positive findings of the AGI study therefore the frustrations deserve consideration.

The AGI sessions were planned, delivered and reflected upon in the tradition of developmental and therapeutic groupwork. The experience of running both groups has left the facilitator with the impression that the AGI sessions were better received by

children and more engaging in the range and nature of the activities offered. The activities themselves were tailored to the group's interests, developmental stage and ability. The processes were the subject of practice supervision sessions conducted between the therapist and a senior work colleague on a monthly basis. The CSG sessions were conducted in a humane child-centred fashion: work was tailored to the group's learning needs and rewards in the form of feedback and class merits were abundant but appropriate and focussed clearly on desired achievement in task or behaviour. Efforts were made to avoid the more global or general positive feedback that was thought to be less helpful in encouraging change, it being more diffuse and having an effect of devaluing the currency of praise. Nevertheless it is inescapable that the AGI sessions were designed to be a richer experience for members with an explicit focus on the development of emotional and social skills through a brief Intervention. Is it therefore reasonable to expect more consistency in the reporting of positive findings?

To be clear: Hypothesis 3 looked for differential gains over the Intervention period and was split, with the teachers and parents finding difference and the children not. The non-questionnaire data (Hypothesis 7) did not report any difference. Only the teachers in Hypothesis 8 are able to report greater gains for the AGI group over the follow-up period, and none of the other respondent groups can distinguish between AGI and CSG at this point. The most simple and in some ways the most persuasive explanation is that the effects of being in a group are powerful but the content of the group sessions does not matter. It could be true to say that the extent of planning for sessions, account of individual needs and session content is of no consequence provided certain core group conditions are met. Yet if this is true, it is perplexing, and the fact that this proposition is both counter-intuitive and counter to received notions of good practice means that it is worth another look. Is it possible to identify extraneous effects to be excluded from any future study?

Over the intervention period, teachers and parents report an advantage of being the AGI group. There is a question mark over whether these teachers are subject to bias, because by and large they are not blind to Intervention status, and most of them will be aware of the relative investment of the therapist in the intervention delivery favouring the AGI group. This may well make them more willing to look for gains in the AGI group as these children are subject to more scrutiny over the intervention period. Countering this

bias is the professionalism of the teachers, and the fact that not all teachers were sympathetic to the research aims and some resented the children being withdrawn from class. In addition, the positive findings reported at T4, where teachers are blind to treatment status, support the positive findings at T3. If these positive and negative biases are judged to cancel each other out and the teacher findings are judged to be valid and reliable, there is still the question of why the findings are not echoed more consistently by the other respondent groups. The parents support the teachers over T2-T3 in distinguishing between the groups, and this finding highlights the effect on emotional problems that the AGI has over the CSG. This is in keeping with expectations from the nature of the two conditions. These parental findings have washed out by T4 and are not present at all in the child data.

This is the most exciting finding from the data relating to the question of whether the groupwork interventions can be distinguished, namely the dual discrimination between the AGI and CSG that is made by teachers and parents over the intervention period T2-T3. Teachers distinguish the AGI from CSG on internalising and externalising scales and the parents do so on internalising behaviours. To say that this is a weight of evidence is perhaps a sympathetic reading of the results, but it is highlighted because of the intriguing insights that it offers into the nature of change and the way it is reported across respondents. That differential effects of a school-based intervention are most readily detectable to teachers is unsurprising, that they are detectable to parents is exciting. The parents register only a differential impact on internalising behaviours, not externalising, although it has been suggested in these pages that parents lacking in emotional literacy skills might not be very good at detecting change in emotional states in their children. These cross-contextual changes present in the AGI intervention can be attributed to the nature of the session content and in particular the training for transition exemplified in the re-creation of home-based scenarios through role-play. If the claim that these effects can be thus attributed is true, it may explain some of the difference in reported attitude to family on the MSCS. These differences are statistically significant, and show the AGI group rating family relations as deteriorating and CSG rating them as improving. This is happening at the same time as parents reporting an improvement in the internalising behaviours of the AGI group. An intriguing finding that has received some discussion earlier due to the direction of change being in the opposite direction to that initially

expected. Attention is drawn to it here to show the triangulation between different measures and the various ways change is reported.

The children report strong positive effects detectable at one-year, but do not distinguish between the groups, either over intervention, or at follow-up. The question here is about whether this is a true effect, or whether there is some variable at work that needs to be excluded. If this is a true effect, the implication is that it does not matter to the children what the content of the group sessions is. Children perceive themselves as improving in self-concept and in reduction of emotional and behavioural problems as an effect of being in a group. Now it could be that the difference in the group content exists mainly in the head of the facilitator: that what the facilitator says and what the children hear are two different things. It could be that the active ingredients of the group sessions are communicated unconsciously and the attempts to deliver content based on differential theory are simply overlooked by the children. On the other hand it would be quite difficult for children to mistake which group they were in. They were run to have a very different content flavour: CSG sat around a central desk with pencils and notebooks, while AGI sat in a circle of chairs and were engaged actively. It would be difficult to confuse the two interventions, but the more active AGI could have been misconstrued as a bit of a lark, which would have detracted from its effectiveness. If this was the case, then to improve the differential effects the AGI should be more focussed in terms of outcomes made explicit to the children and run with a task orientation. It is reasonable to assume that this focus would not have an adverse effect, because the CSG sessions were run in this way.

The alternative to the lack of differential effect reported by children being a true effect is the possibility that some extraneous variable is at work in the child reports. It could be, for example that the CSG group has a high self-concept based on shallow principles. This argument was given some airing in Chapter 4 (it could equally be that the AGI has a false self-concept, but that the teachers and parents are reporting behavioural change). A false self-concept would be characterised by unrealistic appraisal of one's own problem states. Does the CSG with its praise and affirming feedback make children feel better without challenging the basis of their behaviours? This would mean that the CSG children could get by in school because they were boosted in curriculum skills and the school assessment is likely to be influenced by curriculum performance, but might

struggle transferring behaviours outside of the situation they received training in. The parents report consistently lower mean scores for the CSG group over intervention: (T2-T3, AGI= 54.47 – 50.39, CSG= 54.77 – 52.15), which could account for the lack of effect reported by parents for the groups considered together over T2-T3 (the higher CSG scores cancelling out the lower AGI scores). There are signs that the CSG group has made gains that are less stable over contexts than the AGI group, and that the child data are subject to examination of the basis of the shallow nature of the gains self-reported. One reason for this could be the explicit training for transition that took place in the AGI sessions (see section 3.2.2), and it is reasonable to expect that enhancing this feature could result in increased cross-contextual gains.

6.2.2: Extraneous Effects Impacting on Data

Before synthesis is attempted, there are influences that need an account before the conclusions from the data can be stated with confidence. These influences are: regression to the mean effects and permanent exclusion of extreme cases by school. These phenomena have the potential to impact on the findings in a way independent of the intervention and fall outside the remit of the control measures adopted in the AGI design.

Regression to the mean effects:

Regression to the mean effects are found when, as a natural consequence of the passage of time, low scores increase and high scores decrease. In traditional experimental designs, a no intervention control gives an account of regression to the mean effects by monitoring the passage of time in a dummy intervention. There is no no-intervention control in the AGI study and the waiting list condition, while it provided data that was interesting (especially in the case of the YSR and MSCS), did not allow the T1-T2 to act as a no-intervention control, measuring the passage of time as anticipated. Regression to the mean effects would act on the AGI study dataset in a way that would reduce the high scores of disaffected children through a mechanism not due to intervention. For these reasons, some post hoc account of regression to the mean effects is important, particularly as the sample has targeted children at-risk of disaffection and therefore likely to score highly on the measures used and likely therefore to be subject to regression to mean effects.

A measure of regression to the mean effects (R2M) over a given period is the correlation between the initial scores and the gain scores for the period. This is because the nature of R2M is the unequal way it affects low scores and high scores, both moving toward the mean score for the population over time. If this is predominantly the characteristic of the gains made over a given period, then the scores closer to the mean will be less effected by the gains expressed for the population as a whole, and these scores will correlate closely at the beginning and end of the time period. If on the other hand, every case shows an improvement of 3 (for example) then the correlation will be zero indicating no R2M effect. With this in mind it is possible to make some predictions to test in post-hoc analysis for the experimental sample, and these will centre on whether the intervention makes a difference to the nature of changes reported by the populations (AGI, CSG, AGI+CSG). This will be done by comparing the intervention period T2-T3 with the waiting list period (T1-T2) and the follow-up period (T3-T4), to see if the correlations between the gain scores for the period and the initial period reveal patterns other than what might be expected in R2M effects (characterised by a high correlation). An analysis program called RCOMPARE was used to test the significance of the difference between the two correlations in question (Downing 2004). This program provides a z-score which is defined as the difference divided by the standard error of the difference, and the z-value is related to the significance level in as far as data with a value of $z=1.96$ can be assumed to be statistically significant at the 5% level of confidence ($p<0.05$) and data with a value of $z=1.64$ can be assumed to be statistically significant at the 10% level ($p<0.1$).

The first prediction is for the whole group (AGI+CSG), and the prediction is for an effect of intervention over the effect of regression to the mean. This can be inferred for the experimental sample from a difference between the correlations for the waiting list period and the intervention period. The follow-up period is predicted to continue the effects of intervention and show no difference from the intervention period. It is hard to predict if the interventions will combine to enhance or reverse the gains characterised by R2M effects and the prediction does not include a direction of difference. Prediction A: There will be a difference in the correlation between gain scores and initial scores in the period A:T1-T2 and the period B:T2-T3 and no difference between T2-T3 and period C:T3-T4. The data from TRF and YSR support the prediction in showing statistically significant difference between periods A and B but not between periods B and C. The data from

the CBC and MSCS do not show any statistical significance between any periods. Overall the support for the prediction is qualified. This is interpreted as showing signs that the intervention has an effect over and above the naturally occurring effect of R2M. The intervention emphasises the R2M effects by impacting upon cases furthest from the mean.

The second prediction is for the AGI group, and the prediction is for there to be an effect of intervention with those scoring most highly showing the greatest gains as a result of being targeted during AGI sessions. In addition, highly conforming children will be encouraged to express themselves and draw nearer to the mean. These effects will continue after intervention. Prediction B: Correlations between gain scores and initial scores in the period A will be lower than similar correlations in the period B, but there will be no difference between periods B and C. The data from the TRF and MSCS support the prediction in showing statistically significant increase in correlation between periods A and B but not B and C. The data from the CBC and YSR do not show any difference between any periods.

The third prediction is for the CSG group, and the prediction is for there to be no effect of intervention on correlations, with those scoring in both tail quartiles of the distribution curve showing gains of similar magnitude to those closer to the mean as the curriculum intervention uses targeted but equal praise and adult attention uniformly to increase academic competence. Prediction C: Correlations between gain scores and initial scores will show no difference between periods A, B & C. The data from CBC, TRF and YSR support this prediction.

Table 6.2.2a: Correlation coefficients between gain score and initial score and significance value (two-tailed)

		A: T1-T2		B: T2-T3		C: T3-T4		A vs B		B vs C	
		crrte	p =	crrte	p =	crrte	p =	Z=	p	Z=	p
CBC	Whole	0.29	0.09	0.37	0.01	0.44	0.00	0.39	>0.1	0.49	>0.1
	AGI	0.02	0.96	0.36	0.03	0.43	0.02	-1.09	>0.1	-0.31	>0.1
	CSG	0.41	0.07	0.38	0.02	0.42	0.02	0.11	>0.1	-0.18	>0.1
TRF	Whole	0.17	0.12	0.50	0.00	0.51	0.00	-2.56*	<0.05	-0.08	>0.1
	AGI	0.26	0.12	0.56	0.00	0.51	0.00	-1.77*	<0.1	-1.34	>0.1
	CSG	0.14	0.37	0.40	0.00	0.50	0.00	0.39	>0.1	-0.66	>0.1
YSR	Whole	0.36	0.01*	0.58	0.00	0.50	0.00	-1.61*	<0.1	0.71	>0.1
	AGI	0.22	0.26	0.52	0.00	0.45	0.00	-1.42	>0.1	0.43	>0.1
	CSG	0.52	0.01*	0.65	0.00	0.59	0.00	0.72	>0.1	0.43	>0.1
MSC	Whole	0.25	0.15	0.36	0.01	0.54	0.00	0.62	>0.1	-1.39	>0.1
	AGI	-0.14	0.56	0.34	0.03	0.56	0.00	-1.69*	<0.1	-1.20	>0.1
	CSG	0.49	0.06	0.43	0.01	0.51	0.00	0.24	>0.1	-0.43	>0.1

Reflection: overall, the characteristic of the dataset is to show higher correlations between gain score and initial score for all intervention populations over the intervention period T2-T3. These increased correlations are significant in a way that allows qualified support for the predictions, but the pattern is for a definite effect of intervention that favours those scoring in the quartile range furthest from the mean, i.e. those most disaffected. On the basis of these findings, it is possible to say with qualified confidence that the gains described in the AGI study are independent of regression to mean effects.

The exclusion of extreme cases by attrition:

This effect is an effect not attributable to intervention but rather an effect of attrition that could be disproportionately selective to those children most disaffected. School systems routinely exclude (temporarily and permanently) pupils thought to be exceeding boundaries. In addition the most disaffected children themselves are least likely to conform to processes such as those of data collection and most likely to drop out of the study. Because the effect of these exclusions and dropouts could disproportionately effect higher scoring individuals, there could be a significant positive effect on the data from the AGI study that is not attributable to intervention. In order to account for this possible effect, an attempt is made here to quantify it.

This was done by looking at the frequency statistics for CBC, TRF, YSR and MSCS at T2 (T2 favoured over T1 because of the larger number of cases). These statistics describe the interquartile ranges for the instruments and give a cut point for the most extreme quarter. This value was used to select cases from the study population and this selection was examined to determine which cases present at T2 were present at T4. The process was repeated using the cut point to select those in the less extreme quartile ranges. These values were used to determine whether the proportion that dropped out in the extreme quartile was significantly different to the proportion that dropped out in the non-extreme quartiles. This was done using a standard formula (Chambers 1964) section titled: 'significance of difference between proportions'. Guided by this method, the difference in percentage attrition was calculated, and a simple analysis used to find out if the difference in proportion was statistically significant. This was done by dividing the difference by the standard error of the difference to derive a z-score, from which the p value can be derived. Statistical tables show that at z score of 1.96 corresponds to

p=0.05 (the 5% level) and a z score of 1.64 corresponds to 0.1 (the 10% level). The standard error is calculated in the formula:

$$se= \sqrt{pq \times (1/N_1 + 1/N_2)}$$

...where p is the total average and q = (1 – p). N₁ is the number in the first population and N₂ the number in the second. Numbers for AGI and CSG do not necessarily equal the numbers for whole population because the distribution of scores is different in each of those three populations.

Table 6.2.2b: Attrition levels according to severity; numbers and analysis

		Extreme (1 quartile)			Non- extreme (3 quartiles)			z	P <
		Cut Pnt >	T2 n	T4 n	Cut Pnt <	T2 n	T4 n		
CBC	whole	63	25	15	63	66	48	1.176	0.1
	AGI	63	11	6	63	32	22	1.169	0.1
	CSG	65	14	9	65	34	26	1.156	0.1
TRF	whole	63	38	34	63	82	71	0.130	0.1
	AGI	66	14	12	66	44	39	0.500	0.1
	CSG	63	13	12	63	49	42	0.420	0.1
YSR	whole	62	26	23	62	75	60	0.934	0.1
	AGI	63	12	11	63	37	29	-1.030	0.1
	CSG	61	10	9	61	42	34	-0.677	0.1
MSC	whole	58	25	20	58	72	61	0.547	0.1
	AGI	57	11	8	57	36	30	0.779	0.1
	CSG	59	9	7	59	41	36	0.787	0.1

Reflection: the analysis of the data do not give any evidence of a statistical difference between the proportions dropping out of the extreme quartile and the proportion dropping out of the non-extreme quartiles. These two post-hoc analyses of the regression to the mean and the extreme cases reveal that the dataset is not significantly affected by these two extraneous factors and that the findings reported can be treated with some confidence when reporting change and speculating on the nature and causes of change.

The Effect of Administering a Waiting List Assessment

This section examines the extraordinary finding of significant self-reported gains in the waiting list period (T1-T2), when no groupwork intervention took place, as measured by the YSR and MSCS. This finding from the children is not echoed by the parents or

teachers, who report no change in the period T1-T2 followed by significant change over T2-T3, although this is limited to a marginal level of statistical significance in externalising behaviours as reported by parents (together with the effect size findings, this was judged to be sufficient evidence to accept the experimental hypothesis 1d). Explanations for this disagreement between informants were offered in section 5.1, and although there can be some certainty about there being an extraneous factor impacting on the self-reported data, the explanations offered are tentative. These explanations settled on so-called selection and conformance effects that were proposed to arise out of an anticipation of a benign (and attractive) intervention offered as part of the school process. These effects were not detectable in parent and teacher reports. In order to move towards some clarity on the matter, a post-hoc hypothesis is offered. This hypothesis is made possible by the fact that the dataset contains a proportion of children (about half in YSR, MSCS and CBC, a quarter in TRF) who did not receive the waiting list assessment. The hypothesis predicts there to be a difference between the pattern of gains shown by the group of children who did not have the waiting list assessment compared to those who did. In accordance with the selection and conformance effects it is predicted that the children that had no assessment at T1 will show significant gains over period T2-T3, while the children who did have the T1 assessment will show significant gains over T1-T2, and no gains over T2-T3. A sub-hypothesis is offered because of the expectation that the self-reports will differ from the teacher and parent reports based on the results from hypothesis 1. This post-hoc hypothesis can be stated thus: children with no assessment at T1 will show gains over a different time period than those assessed at T1. The sub-hypothesis can also be stated: the pattern of gains shown by analysis of the self-reports will be different to those shown by the parent and teacher reports. Total T-scores for YSR, MSCS, TRF and CBC were used in a paired samples T-test, testing T1_{ts}-T2_{ts}, T2_{ts}-T3_{ts} and T3_{ts}-T4_{ts} for each instrument. The findings are presented in the following table.

Table 6.2.2c: Change over time: pattern of gains made by those with waiting list assessment (T1=1) and those without (T1=0)

	T1 –T2		T2-T3		T3-T4	
	P<	n	P<	n	P<	n
pres Y1=1	0.01 (0.001)	52	ns (0.117)	45	0.1 (0.083)	40
pres Y1=0	N/A	-	0.05 (0.42)	39	ns (0.624)	41
pres M1=1	0.01 (0.002)	35	ns (0.472)	30	ns (0.244)	27
pres M1=0	N/A	-	0.01 (0.003)	51	0.05 (0.037)	53
pres T1=1	ns 0.345	82	0.01 (0.000)	82	0.05 (0.025)	71
pres T1=0	N/A	-	ns (0.512)	31	ns (0.357)	29
pres C1=1	ns (0.961)	35	0.1 (0.056)	31	ns (0.459)	24
pres C1=0	N/A	-	0.01 (0.000)	45	0.05 (0.041)	38

Table 6.2.2d. Mean scores of those with waiting list assessments (T1=1) and those without (T1=0)

	TOTAL T-SCORE MEAN SCORE	
	Pres T1=1	Pres T1=0
YSR T1	59.50	N/A
YSR T2	55.31	52.64
YSR T3	52.45	48.46
YSR T4	49.10	49.61
MSCS T1	54.57	N/A
MSCS T2	50.07	53.24
MSCS T3	48.52	49.70
MSCS T4	46.37	47.43
CBC T1	56.94	N/A
CBC T2	56.13	53.40
CBC T3	52.04	50.63
CBC T4	53.33	53.82
TRF T1	58.60	N/A
TRF T2	57.98	54.61
TRF T3	52.55	53.10
TRF T4	49.08	51.14

The analyses conducted give split support to the post-hoc hypothesis, the YSR and MSCS behaving as predicted and no support given by the CBC and TRF. The YSR and MSCS when examined in Hypotheses 1a and 1b showed gains in the waiting list period, and when examined here, children with a waiting list assessment show statistically significant gains over the waiting list period, followed by no change over the intervention period T2-T3. Those without the waiting list assessment show statistically significant gains over the intervention period T2-T3. The parents, when reporting on change observed in their children do not distinguish between gains made T2-T3, and although the teacher data are interesting, they do not report gains by the non-assessed children over any time period. The split support given to the post-hoc hypothesis is understandable in the light of the fulfilled predictions of the sub-hypothesis, that the expectation of a difference between the self reports and other reports is justified when the findings from other hypotheses is taken into account.

Reflection: There is a striking difference between the children who had an assessment (As1) and those who did not (As0) when self-reports are considered. The findings do seem to support extraneous effects such as a 'selection effect' or a 'conformance effect', whereby there is an immediate positive impact on problem behaviours and self-concept following the administration of the assessment tools. Speculation about this effect could reasonably focus on the child's anticipation of an attractive event that is perceived (even when in distant prospect) to provide pleasant variety to the school day. Caution is needed, however, because the As0 children show a significant improvement over T2-T3, where extraneous effects would coincide with effects of intervention, without the possibility of distinguishing between the two. It is also interesting that the As0 children rate themselves with less severe behaviour problems the first time they are assessed, this is relevant to the reflection offered on the TRF. The parents show As1 with no change over T1-T2, followed by both As1 and As0 significantly improving over T2-T3 and both declining over T3-T4, As0 with statistical significance. This corresponds to the patterns observed in discussion of Hypotheses 1 and 2, and is echoed by the YSR scores for As0 above, where a slight decline in scores is observed T3-T4. These reflections illuminate findings from the TRF: here the As0 group show no significant gains over T2-T3 or T3-T4, while the As1 group improve with statistical significance over both time periods. It is possible to say that children assess themselves to be benefiting from the experience of being assessed and part of an intervention, and the

excitement/anxiety generated leads them to show gains over the period of the experience not necessarily related to the period of intervention. Those children who receive an intervention soon after initial assessment reveal themselves to have less problematic behaviours and the behaviour gains do not hold up so well over time. This pattern is strongly echoed by the parents and teachers, both of whom assess the As0 group to be less problematic at initial assessment, and both parent and teachers show the As0 group to be less able to sustain gains over time. The MSCS measures self-concept and this appears to show a different trend for the As0 group, showing them to continue improving over time (T3-T4) in a way that exceeds the As1 group. The explanation for the apparent tendency of the As0 group to score more positively at first assessment may be due to methodology in as far as those without the T1 assessment were the schools in the first wave of intervention and the communications with schools were more likely to have been rushed and messages (without the advantage of practice) more likely to have been mixed. In this way, the possibility of misunderstanding the selection criteria would be increased and relationships with classroom teachers insufficiently established with the effect of withdrawing children generating resentment, reflected in a lack of sympathy with the aims of the AGI study. If it was the case that selection criteria had not been communicated adequately, this would explain the finding from the parents that As0 children were assessed as less problematic at T2 than As1 children were at T1. Also it would explain why improvement in self-concept (MSCS) in the As0 group is at a high level of statistical significance, while As1 gains are not significant over any time period, it is reasonable to expect that children with less problematic behaviour would find themselves more able to build and sustain positive self-concept. Overall, it is difficult to conclude otherwise than by saying that using the population who did not have the waiting list assessment to illuminate the counter-intuitive gains reported on YSR and MSCS measures over the waiting list period has mixed success. The reflections it has facilitated on methodology have been useful, but the impact on the establishment of proposed selection and conformance effects has been reduced by the reflections. It is nevertheless to accept with qualified confidence the existence of such effects as an explanation of the YSR and MSCS gains over T1-T2.

6.2.3: Synthesis

The patterns in the data have been given the explanations in section 5.2 above with the aim of adding to the interpretations of individual hypotheses given in Section 5.1. The process of generalisation from individual hypotheses to patterns connecting all

hypotheses and the broadening of explanations from the specific to those able to incorporate the connecting patterns allows a synthesis to be made at this point. In other words, the data, the findings and the explanations offered to understand the findings all make the case for an effect of intervention that can be synthesised below.

There is robust and multi-source data showing that brief group interventions can reduce youth disaffection, and these interventions make a difference over and above the usual input of the school environment. The beneficial effects of this groupwork last up to one year and are felt by class teachers and the children themselves. The children report a reduction in emotional and behavioural problems and an increase in positive self-image. The teachers report a reduction in emotional and behavioural problems. These benefits of intervention are sustained one year after the intervention finishes. In addition there is evidence that the time spent on-task in curriculum time increases as a result of the intervention.

Differential effects favouring the AGI condition are found over intervention by teachers and parents and by teachers at one-year post-intervention. This provides evidence of the superior effectiveness of the AGI intervention in the short-term, delivering differential benefits in the school and transferring out to the home. In addition the AGI is found to be more effective in promoting long-term gains in the school.

These results reveal a general picture of gains more easily felt in the situation they were delivered in, and lasting longer in school than in the home. Cross-contextual gains were found for the AGI condition T2-T3, where the intervention specifically trained children for home-school transfer and this may explain some of the difference in reported attitude to family on the MSCS scale. This allows a socially constructed reading of disaffection, but highlights the need for explicit reference to situations over general principle.

This synthesis reveals powerful interventions effective in meeting the challenges of disaffection: scale, complexity and need for outcome. The challenge of scale has been met by the development of brief group interventions that are adaptable and straightforward to deliver. The challenge of complexity has been met by session content aimed at developing relevant skills and positive internal states. The challenge of

outcome has been met through the rigour of the research process, the multiple respondents and the positive outcome in a number of areas of child function.

Section 6.3. Discussion

6.3.1: This study in the context of other studies

There are significant successes to be taken from the findings interpreted above, but what is the measure of this success? There have also been significant areas where expectations have been frustrated, but to what extent have these been anticipated in the literature? The focus of this section is to use the literature to find a relative position from which to regard the successes and limitations of the study.

Positive findings are not unusual in the literature of child and adolescent disturbance, but as discussed in chapter one, findings reporting on group interventions in a randomised control design taking place in a community setting with longitudinal follow-up are much rarer. This makes the task of referencing the current project with exact matches for design a challenging task, but not to do so runs the risk of not comparing like with like and so compromising the conclusions of relative impact. For example, comparing the impact of the interventions in this study with studies focussing on a depressed (e.g.) child population rather than one of mixed symptom/multiple diagnostic category is only a partial intersection with the TLC study and is only a partial comparison.

Importantly for this discussion, studies that contain data comparing multiple intervention conditions are less usual. This is important because the predictions frustrated by the data centre on their inability to distinguish between the group interventions. For the purposes of this discussion this design feature is given higher prominence, because it is this aspect of the data that is most challenging to explain.

The table below uses the literature reviewed in Chapter One to identify the common design features between the current study and the published studies. This is done with the aim of clarifying the areas of intersection between studies and in order to justify the attention given to individual studies in the following discussion.

Table 6.3.1:
Match between Design Components of Key Studies and those Identified as Desirable

	AGI 2004	Kolvin 1981	Cowen 1996	Weiss 1999	CPRG 2002	Reynolds 2003
School Context	Yes	Yes	Yes	Yes	Yes	Yes
Randomised Control	Yes	Yes	No	Yes	Yes	Yes
Longitudinal Design	Yes	Yes	Yes	No	Yes	Yes
Multiple Respondents	Yes	Yes	Yes	Yes	Yes	Yes
Real-World Outcomes	Yes	Yes	Yes	No	Yes	Yes
Multiple Intervention Conditions	Yes	Yes	No	Yes	Yes	Yes
Time-limited (brief) Intervention	Yes	Yes	No	Yes	No	No
Group Interventions	Yes	Yes	No	Yes	Yes	Yes
Targeted Prevention	Yes	Yes	Yes	Yes	Yes	Yes
Multiple Diagnostic Category	Yes	Yes	Yes	Yes	No	Yes
Positive Parametric Data: p<.05	Yes	Yes	No	Yes	Yes	Yes
Differential Group gains	Yes	No	No	No	Yes	No

Comparing the current study with published studies that do not contain a certain number of common elements is counter-productive. Using this table it is possible to decide which studies make a good fit with the current project and a number of these are discussed below, This is done with the aim of providing a reference point for both the success of the current project (where is has made a lasting impact) and the limitations of the project (where predictions have been frustrated).

For the sake of clarity these areas of success and limitation are listed below.

Success:

- Positive change in pupil self-perception relating to problem behaviours
- Positive change in pupil self-perception relating to self-image
- Positive change in teacher perception relating to problem behaviours
- Positive impact on classroom engagement measured by time spent on-task
- The Increase in rate of change at start of intervention compared to waiting list no-intervention reported by teachers and parents
- Sustaining positive change in school at one-year follow-up
- Sustaining positive change in self-perception at one-year follow-up
- Short-term differential gains between interventions reported in school

Short-term differential gains between interventions reported at home

Long-term differential gains reported in school

Limitation:

Perplexing self-reported gains over waiting list period

No self-reported differential gains

Washing out of gains in the home over time

The lack of positive data relating to classroom disruption

The frustration of predictions about attendance

The study that bears most resemblance to the current project is that conducted by Kolvin and published in 1981. This study assessed the impact of interventions on discrete populations of junior schoolchildren (primary school) and senior schoolchildren (secondary schools). Each population was given a parent intervention or a group therapy intervention or a behaviour modification intervention. These interventions were compared to each other and to an at-risk control group over intervention and at follow-up points up to midline (12 months post intervention), and final (36 months post intervention).

Comparisons here with Kolvin concentrate on the senior school children as the closest match in terms of age for this study's sample.

The Children

TLC: children report significant improvement over the intervention period and sustained at one-year follow-up on both problem behaviours and self-image.

Kolvin: The study uses an aggregated measure for problem behaviours and what he calls 'neurotic anxiety'. These will serve to compare with the problem behaviour and self-image data respectively as used in the TLC study. Kolvin's data show children's self-assessed problem behaviour getting worse over intervention and at midline and at follow-up, while their 'neurotic anxiety' improves to midline and shows no significance at follow-up.

Reflection: The gains reported by the children in the TLC study appear to be robust and deep-rooted.

The Teachers

TLC: teachers report significant change over intervention as compared to waiting list and sustained improvement at follow-up.

Kolvin: teachers using Rutter Scale B2 report no change to midline and significant improvement at final follow-up.

CPRG: Children were assessed yearly and results reporting at 3 years show an effect of intervention using an amalgam including the Externalising scale from the TRF. Two variables yielded significant main effects for intervention, but there was no effect on the TRF.

Reflection: A feature of the Kolvin study is the so-called 'sleepers effect' whereby changes not present at one year (midline) come through at 3 Years. It gives cause to wonder what the effects at 3 years will be for the TLC cohorts. The CPRG data shows teachers inconsistently reporting change due to intervention at 3 years.

Curriculum Engagement

TLC: shows on-task behaviour significantly improving over intervention compared to waiting list and sustained at follow-up.

Kolvin: shows children's self-reported interest in schoolwork and their relationship with teacher deteriorating both at midline and significantly at follow-up.

CPRG: data on academic progress did not yield significant effects of intervention.

Reflection: To look for and detect benefits of group interventions in curriculum engagement is unusual.

The Parents

TLC: the parent data reports children improving over intervention relative to a waiting-list period, but report no long-term effects.

Kolvin: Using Rutter Scale A parents report significant improvement over intervention and at midline, but no significant change from baseline to final follow-up.

CPRG: two of the four measures yielded an effect of intervention, but both were on derived self-assessment measures. Questionnaire data did not show an effect of intervention.

Reflection: The parent data from the TLC study consistently frustrated the experimental hypotheses and the explanations offered focussed on the inappropriateness of the instrument, the difficulty of cross-contextual generalisation and the lack of recognition given to positive behaviour gains in the home. The Kolvin study helps clarify these

explanations in that it shows gains are possible and that there is a potential for cross-contextual gains to register in the home environment. That having been said, the Kolvin study included a parent intervention, designed to explain and arbitrate school aims across to the home. It could well be that gains in the home will only appear after such an intervention. Reflecting on the TLC findings in the light of the Kolvin study therefore affirms the possibility of the explanations being correct.

Classroom Behaviour

TLC: data collected in the classroom failed to show any improvement in the children compared to the rest of the class.

Kolvin: has no comparable data.

CPRG: measured sociometry but did not report an effect of intervention.

Reflection: the paucity of comparable data from selected studies makes it difficult to generalise from these findings.

Attendance

TLC: attendance falls over intervention and decline accelerates over follow-up

Kolvin: has no comparable data

Reflection: the paucity of comparable data from selected studies makes it difficult to generalise from these findings.

Distinguishing between Intervention Groups

TLC: children reported no differential gains between intervention condition. Teachers reported difference over intervention and at follow-up. Parents reported difference over intervention in internalising behaviours.

Kolvin: data provided on 2 group interventions; Behaviour Modification (BM) and Group Therapy (SG). Data from children are unable to distinguish between the two at any point (p371) and teachers have the same inability to distinguish. Interestingly, parents put SG above BM on the Rutter A scale at midline, but do not distinguish otherwise.

CPRG: parents report a difference between interventions and control at 3 years, as do teachers.

Reflection: The Kolvin study is larger in scale and ambition, giving between 60 and 80 in each intervention condition, close to double the numbers at final follow-up in the TLC study. The inability of this larger study to distinguish between group interventions makes

it harder to support the explanation offered in the TLC study that the sample size is the cause of insignificant findings. Kolvin offers no reflection on this in his discussion, but it does seem to be a feature of published research that data are supplied in support of single interventions, but there is a difficulty in separating the effectiveness of two interventions intended to have a different theoretical base.

6.3.2: Implications for Theory

The position of the TLC findings relative to published studies with common design features is fairly strong: the methodological rigour and community context of the intervention are unusual features, and the findings compare well to published studies. As a result there are a number of areas where the findings from the TLC study, seen in relation to other published studies, can be used to reflect critically on theory.

The positive tone of the findings and the compact nature of the intervention make therapeutic intervention possible in the school setting. A case can be made for enlarging the framework of educational policy and practice to include explicit developmental theory.

The data gathered from the home demonstrates common difficulties with engendering cross-contextual gains, but where present they serve to illustrate a need to explicitly train children to transfer appropriate social skills from the school to the home environment. These findings can be used to usefully reflect on the socially constructed and context-specific nature of disaffection, informing effective intervention in home-school issues. TLC bucks the trend in distinguishing between the intervention groups, but there does seem to be a group effect that is as strong as a content effect. The active factors are ill defined in theoretical and outcome literature

Implications for Educational Policy and Practice.

In Chapter One (section 1.2.3), the case was made for the reluctance of researchers to seriously engage with the potential of the school environment was introduced.

“Schools have long been identified as an ideal entry point for improving access to mental health services for children (Adelman & Taylor 2000) and are set to play a key role in the mental health of young children (The Children Act 2004) and their unique socialising

potential has been recognised by researchers as well as teachers (Kolvin, Garside, Nicol, Macmillan, Wolstenholme, & Leitch 1981), (Brint, Contreras, & Matthews 2001), (Zsolnai 2002). Although there is wide recognition that school-based approaches to the management of delinquency have considerable potential, recent reviews of the literature typically conclude that 'studies to date have failed to demonstrate powerful positive effects' (Fonagy, Target, Cottrell, Phillips, & Kurtz 2002) (p386), perhaps due to the perception that '...few programs have been rigorously evaluated in the real-world setting of schools' (Hoagwood 2000). However, school-based interventions do overcome the disadvantage of clinic-based interventions (in that gains made in clinic are notoriously difficult to generalise) by seeking gains in the locus (or one of the main loci) of disturbance, and the school environment with its ecological setting combined with a relatively constant population, does provide a natural focus for different agencies affected by anti-social behaviour (Adelman & Taylor 2000)."

The knock-on effect of a lack of research into school-based intervention programmes is a lack of evaluated innovation and a consequent confusion about the legitimate scope of school-based interventions in youth disturbance and disaffection. The AGI study adds to those studies able to make the case for effective interventions in the school setting, and it is the combined weight of these studies that makes a strong case for the parameters of school function to be enlarged to accommodate a range of therapeutic services. Therapeutic services such as those detailed in the AGI study can reduce self-reported emotional and behavioural problems and lead to improvements in the home, with the effect of impacting on classroom behaviours.

The implications for the framework of schooling go further when one considers that children's school performance (understood as an amalgam of attendance, achievement and engagement with the learning and social opportunities of schooling) improves when their internal profile improves. Including therapeutic services because they have the potential to improve achievement admits the case for educating the whole child and improving school performance – widening the remit of education itself. If the case for therapeutic services is admitted, the case for complementing the current school emphasis on academic achievement with an equal emphasis on the child as person cannot be avoided.

Implications for the Understanding of Disaffection:

The intervention looked for real-world outcomes, but looked for them as a result of internal changes in developmental psychopathology. The assumption was that these internal changes would translate, in the manner of an internal tool-kit, to situations outside the intervention context. Instead what appears to have happened is that where internal changes (in self-image and self-assessed emotional and behavioural profile) have occurred, the resulting benefit is restricted to the school context, with the conclusion that a school-based intervention allows children to better navigate a school system.

What appears to have been given incomplete account is the socially constructed nature of disaffection, that is to say the way people within specific environments define certain behaviours as problematic and then create categories for dysfunction to contain them. The school definitions of 'problematic' are not the same as those in the home environments of many families, principally because of the socio-economic status of households and the more middle-class norms of the school environment. This being the case, what is problematic for the school is not necessarily problematic in the home. If the child is perceived to be disaffected in school but not at home, can the child be fairly categorised as disaffected? It is true that many disaffected behaviours are challenging across contexts, causing pain and upset to all involved in a way that is not relative to the environment, but it is also true to say that disaffection is largely a social phenomenon, produced by interactions between people within a social environment. Intervening as if the child is the only player in this social environment is misleading. This is the socially constricted nature of disaffection.

The findings from the TLC study in accordance with other published studies show that benefits tend to be context specific. A social constructivist reading of this finding is that this school intervention is delivered in a middle-class environment, by a well-meaning middle-class practitioner, and equips children to navigate a middle class environment. In other words the benefit is social, rather than internal. As a consequence, the learned social skills cannot be expected to translate into the different home environment, where behavioural expectations are different, even when one disregards the socio-economic status of households.

The evidence from the TLC study supports this reading of the impact on theory. As a rule, insufficient attention is given to the socially constructed model of disaffection, with greater attention given to the internal state model.

Implications for What Works for Whom?

The case was made forcibly in Chapter One for an intervention of a particular format (group and community-based) and delivering a particular content (pan-theoretical and developmental in focus). While both the AGI and CSG interventions conform to the criteria of format, the predictions made in the hypotheses indicate that the criteria of content would be more fully fulfilled by the AGI intervention. The findings from this study and others mentioned above tell us that the content of the sessions is of less importance than previously thought. The implications for theory are therefore rather searching.

Interventions are usually allied to a theoretical understanding of children's development, and the form of intervention is prescribed by the interpretation made of children's function. For example, a behavioural understanding of the way children operate will result in an intervention that seeks to redress the perceived imbalances through behavioural means. The forms of intervention are considered as distinct and distinctive as the theories behind them. Chapter One made some study of how the huge majority of the outcome literature focuses on the effect on behaviour of a particular intervention type: "The vast majority of studies will test the effects of a conceptually distinct treatment modality (e.g. CBT) with a diagnostically distinct symptom cluster (e.g. depression), giving a plethora of sub-varieties of treatment and symptom with a competing and confusing set of outcome profiles under which some sort of outcome is usually present for most treatments" (In Section 1.3).

In response to the deficiencies highlighted in such an approach, the AGI intervention was constructed according to emerging and cutting-edge trends in theoretical bases for intervention. "The trends in treatment tend to favour heterogeneous approaches based on integrated understandings of childhood dysfunction proposing systemic change to meet practical ends. These trends provide important legitimising messages to those treating disaffection, where diversity is the starting point and complexity the consequence of targeted prevention. What does one have to do to be effective in

reducing disaffected behaviours? The intuitive rightness of pluralistic approaches is in rare agreement with evidence-based literature and it is relatively simple to extrapolate from these trends to guide effective preventative practice. Overarching and pluralistic theoretical frameworks, and as a consequence multi-method systemic approaches in assessment and treatment will characterise interventions accountable to the challenge of complexity” (in Section 1.3).

In other words the TLC study eschewed the usual modes of comparison, i.e. those between different theoretical standpoints on childhood function. The TLC study opted instead for a pan-theoretical model for intervention and compared this to an intervention based on learning and pedagogical theory – not considered as covering the same areas of child function. But the result is similar to those studies that have such difficulty in detecting difference between (e.g.) behaviour modification groups and group therapy groups (Kolvin 1981).

From this evidence it is difficult to support the proposition that the pan-theoretical approach is the active factor in promoting change. Trends in the literature point to the conclusion that the specific form of theory and intervention (whether you are a cognitive-behaviourist or a social skills practitioner) is not the active factor in promoting change either.

What is the active factor in promoting change remains a crucial question as interventions are planned for this generation of disaffected children. On the basis of the TLC study and the insights gained through placing it in the context of literature on the subject of disaffection, it appears that the active factors will be found in the quality of attention brought to children's experience, whether that experience be their social and emotional lives, or their experience in building up their learning.

Chapter Seven: Conclusion

Introduction

This chapter forms the conclusion of a great deal of work and reading (not least for the reader), and in it an attempt is made to orientate the preceding text and its discussions outward. That is to say, the journey from the incidents described at the beginning of Chapter One through epidemiology and scholarly reflection in the literature of disaffection and intervention to research questions and hypotheses in Chapter Two traced a route towards specific and detailed consideration of the interventions and their evaluation and analysis. In this chapter the journey is re-traced in reverse: moving from the specific to the general. This is done to be clear about what the AGI study has to offer to practitioners wanting to develop systematic practice in dealing with issues of complexity, scale and outcome. The aims of the AGI study are articulated and the discussion of data related to these aims to judge what has preceded this section in terms of what can be learned for practice. In this vein the arc from specific data to practice is continued by considering the varying success of the study in the context of the practice of The Learning Challenge in delivering groupwork in the North east of England. The section on the limitations of the study injects a sense of proportion to the consideration of how much impact this will have on wider practice, in this section the specific issues of reflexivity and bias alluded to throughout the text are engaged. The chapter concludes by considering what the implications for future research and practice might include.

Section 7.1: Summary

This written account of the AGI study has been a practitioner's contribution to the efforts made to meet the challenges of youth disturbance and youth disaffection. By focussing on youth disaffection the intention has been to acknowledge a wider range of behaviours and attitudes as potentially problematic and potentially responsive to intervention. In planning and delivering the interventions, in developing research methodology and gathering data, the intention has been to provide children and chalk-face practitioners with tools to negotiate the issues of disaffection.

The issues have been presented here as issues of scale; as if the anecdotal evidence from the classrooms and clinics were not enough, there is ample evidence that

disturbance and disaffection present a huge challenge to education, health services and the youth justice system. There are also issues of complexity; disaffection is a complex set of attitudes and behaviours triggered by events that cannot be predicted with any certainty, and presents a challenge to traditional treatment frameworks. Finally there are issues of outcome; disaffection is a shifting set of behaviours existing for the most part at lower levels of severity, and presents a challenge to notions of clinical outcome.

Using the literature to establish these issues as the key challenges of disaffection, scale, complexity and outcome were placed at the centre of any intervention that could be considered effective. Using the literature in published studies as a guide, components were presented that an intervention could include. These components were found to be effective in reducing the incidence of disturbed and disaffected behaviours, and able to meet the issues of disaffection. The components were summarised as:

Issues of Scale

A targeted prevention focus

A group intervention approach

A non-clinic, real world setting

Issues of Complexity

A systemic understanding of disaffection (the contributing role of the environment)

Intervention method based on an integrative analysis of disaffection

Content that reflects 'meta-themes' of effectiveness

Issues of Outcome

Rigour in methodology and commitment to outcome

Outcome measures from multiple viewpoints

Awareness of cultural determination in issues of outcome

Published studies were interrogated to establish precedence and to define the contribution of the AGI study. Key studies were identified as exemplars of good practice in their ability to deliver quality defined under the components above. These key studies had the function of benchmarking and measuring the success of the AGI study. In

detailed examination of the methodology and outcomes from these studies, the definition of the AGI study was possible.

The triggers for the research were practical in the first instance: the research team perceived a need for a theory-based intervention able to meet this formulation of the issues of disaffection. The analysis outlined above is implicit in many places in the literature, but presented in the AGI study as explicitly guiding the form of intervention. There was a perceived need for methodological rigour in school-based studies and outcome focussed group interventions. In addition the children were multiply assessed through questionnaire and observation, giving a unique combination of evidence in one study. This built robustness into the assessment and an imperative to any positive findings, for the AGI study grew out of practice and was intended at every stage to be easily incorporated back into practice.

The research questions that gave focus to the perceived need reflected this concern with practice. The AGI team wanted to know what happened to children over time as a result of the interventions offered, whether there was any difference between the interventions and whether there were behavioural changes in the classroom and school. From these research questions hypotheses were developed to test key aspects and impacts of the interventions over time.

The methodology chapter made the case for understanding epistemology and ontology in the light of practice, and proposed an approach that made the best of a variety of methods and research designs. Methodology was formulated containing a mixture of quantitative and qualitative methods and directed to key audiences with the aim of providing persuasive evidence to initiate change where appropriate. This section also made a clear statement of the steps followed and gave a description of the interventions themselves and the population of children that received intervention, before making the case for an analysis protocol that made high demands on the data in a search for significance, both statistical and practical.

Chapter 5 was concerned with results. Initial reporting set out the hypotheses in sub-sections relating to the data-gathering instruments. Data were analysed using ANOVA to give a value for statistical significance, and effect size as a measure of the magnitude of

change. Effect size is particularly useful when sample sizes are small and was found to be decisive in several cases relating to findings from the hypotheses. Presentation of the data was focussed on the support or otherwise of the experimental hypotheses, and the hypotheses all predicted change in one direction. Six of these hypotheses were not supported by the data. Four of the six hypotheses looked for difference between the group intervention conditions. The others related to findings from the waiting-list period. Interpretations made in Chapter 6 placed these findings in the context of key studies identified in chapter one to act as benchmarks and measures of success. The AGI study was found to reflect common difficulties in establishing particular types of gain. These interpretations celebrated the positive tone of the outcome data, and looked for explanations when predictions had been frustrated. The most significant areas where explanations were sought were for the parental data that showed no impact of intervention, and the inability of the data to distinguish between the intervention groups. Published studies were again reviewed to gain insights from precedents and the resulting conclusions used to reflect on theory.

The positioning of the study in rigorous methodology and practical approach is intended to provide a platform on which the findings can be displayed with credibility. This credibility is intended to be two-fold, with implications for practitioners and policymakers outlined below

Section 7.2: Potential Practical Impact

The TLC study set out as a practical project in that it began out of the experience of a practitioner trying to deal effectively with challenging youth. The aim of engaging with the subject in such an extended way has been to support the effective practice of frontline staff dealing with disaffection. The study has maintained that this bottom up approach has influenced all aspects of the research project and the theoretical insights that have both underpinned the process and arisen from the findings.

It is therefore right that the process should be finishing with a reflection on how the work represented here can impact on practitioners interfacing with the challenges of disaffected children daily. For these teachers, classroom assistants, learning mentors, school health workers, community psychiatric nurses, educational psychologists and educational welfare officers (to name a few) the primary concerns are practical ones.

How do I manage my caseload? How do I make an impact to the learning environment in the school? How can I reach the numbers of those disaffected? What do I do with them? How can I be sure that the work I do is effective? How can I engage the unengaged, teach the untaught, break down resistance, bring a taste of the joy of learning to jaded palates? How can I break the cycles of petty violence that characterise the peer group?

To these people the TLC study aims to bring some solutions. What does the TLC study have to offer? First and foremost the broad agreement of data sources in their support of group interventions as tools to reduce the incidence of disaffection in schools is the key tool to release practice. Practitioners can work with groups of children in a way that brings down the caseload and is effective in reducing disaffection. The group setting means that it is possible to work with a section of the school population, rather than with an unending stream of individuals. It means that strategic work is more possible as behaviours are anticipated and challenged before they begin their huge impact on workloads.

The second contribution of the TLC study to the over-worked lives of front-line practitioners is the format of the group provision: these are brief interventions, delivered in school, in school time with an impact that can be felt one year after intervention: this is a powerful, but manoeuvrable, tool. The accessibility of this tool for school staff and the realistic time commitment involved in delivery gives a sense of possibility to the task of making an impact in the lives of disaffected children.

These characteristics mean that it is possible to impact upon school culture. As larger numbers of children reflect the gains of the group intervention, there will be a point where all of the at-risk population will have received this provision. If gains are made and maintained by the group are significant (as this study shows they might be), this will create a powerful effect on the peer group. And school culture is not limited to the children: staff have a huge role in determining the cultural norms, and are under-represented in studies into the mental health of schools (Jackson 2002). A large-scale mental health intervention of the type described above will necessarily impact on the expectations on staff, as support for the intervention is requested from management.

The credibility of the TLC study is intended for policy-makers as well as practitioners. The benefits felt by children and the school are very positive and a significant finding of this study. Positive findings are not unusual in the literature of child and adolescent disturbance, but as discussed in chapter one, findings reporting on group interventions in a randomised control design taking place in a community setting with longitudinal follow-up are much rarer.

The findings from this study are a significant addition to the literature for these reasons of design alone. The design takes seriously issues of disaffection and context, methodological rigour and diversity, outcome and sustainability, and in doing so sets the scene for interventions providing solutions to the behaviours of disaffection, delivered in a way that can be effective in meeting the challenges of scale and context. Through the methodology detailed in Chapter 4 the data are presented as robust enough to stand up to the most testing scrutiny. The robust status that is intended for the data makes the implications for theory more sustainable and therefore the wider practical implications less easy to dismiss. In this category it is possible to see practical implications that might result from the widening of theory that the findings of the TLC study have encouraged. These practical implications sit under the headings discussed in Chapter 6 when theory was considered.

School Provision:

There are implications for the provision of schooling because of the AGI findings. These implications are more difficult to dismiss because of the measures taken to adhere to strict standards for the data. The central implication, as discussed in the last chapter is the importance of an increased emphasis on therapeutic services as part of the normal delivery remit of the school system. The practical implication of this is the need for increased accessibility to therapeutic interventions by non-specialist staff. Therapeutic interventions will have to become more accessible, and not shrouded in so much mystique. To be fully effective as therapeutic institutions, the interventions need to be accessible to those practitioners already active in the school setting. The potential impact of this project is to show that a schoolteacher can deliver a therapeutic intervention that is effective. Moreover he or she can deliver it in a brief solution-focussed design that does not go as deep as traditional therapies, but nevertheless has lasting effects. If the training in this intervention can be shown to be effective in

communicating the active factors of change to others, then the rollout of this intervention becomes a possibility.

Parental Data:

These findings can be used to usefully reflect on the socially constructed nature of disaffection, and the implications for research and intervention practice. The implications of the findings from this study highlight the effect of intervention in the school setting, with the implication that the intervention seems to equip children with the resources to deal better with the hierarchical demands of the school system, but without the same impact in the home situation. The reason for this was thought to be due to the way disaffection could be seen as context-specific, a social rather than internal phenomenon. In the light of this finding, it is possible to predict that interventions that are effective across contexts will incorporate some of the social norms from the contexts they wish to affect. This might be done through explicit role-play, or through sessions conducted outside school time. This is not a new finding, but examples of this formulation of the theory behind the research finding are not as easy to find as might be thought.

Intervention Groups:

The findings from this study and others mentioned above tells us that the content of the sessions is of less importance than previously thought. The implications for this are potentially far-reaching when it comes to deciding what it is that takes place in schools as part of a therapeutic programme. The most convincing of the interpretations offered focus on the communication (or otherwise) of the session content to the children involved. It seems possible that the difference in psychological content favouring the AGI sessions existed primarily in the mind of the intervener and not in the perception of the children. The powerful effect, detectable one-year post intervention, is to be found in the fact of the group rather than its content. This is an important finding for future planning of interventions: it seems that what happens in the groups is less important than the way they are managed. More thought should be put into this aspect than planning content.

Section 7.3: Assessment of Success of Project

The success of the project can be defined as whether it has met its own aims.

Grand Aims:

The AGI study aimed to advance provision in youth disaffection, specifically: advances in the practice of service delivery in school, advances in techniques available to practitioners dealing with the challenge of disaffection, and advances in the behaviour profile of the children who are themselves disaffected. In addition, the aim was to provide advances in the theory of disaffection and advances in policy planning and development in youth disaffection in the school setting. For this reason findings should be suitable for publication in a range of theoretical, professional and research journals. These grand aims could only be achieved if other aims were met:

Pragmatic Aims:

The grand aims were supported by pragmatic aims to develop interventions capable of delivering change in disaffection with lasting effects over time. These pragmatic aims concerned overcoming practical and professional obstacles in working with schools to meeting the challenges of disaffection while building relationships with staff teams and parents.

The aims relating to the development and deliver of interventions capable of delivering change in disaffected behaviours over time have all been met. The intervention is a targeted preventative intervention using the group setting and the school environment to deliver provision of an integrated theoretical base. The interventions meet the criteria of being able to deal with the key issues of disaffection: those of scale, complexity and outcome. These interventions are like hen's teeth. It is a significant achievement to have working with 130 challenging children at-risk of disaffection and to have worked with them in such a way that statistical and practically significant gains have been made in their behaviour that are detectable one year post-intervention. This is testament to an effective approach to group intervention in an area that provides remarkably little good news. In order to achieve this good working relationships with classteacher, headteachers and classroom assistants have been established and environments broadly supportive to the therapeutic aims of the interventions were maintained over the period of involvement with the school. In the light of this, it is possible to say that the pragmatic aims of the AGI study have emphatically achieved.

Research Aims:

In addition there were research aims concerning the development of a methodology able to provide data that were relevant to the intervention and reliable in measuring any change over time. The data generated enable in-depth reflection on the findings in relation to other studies of similar aim. This is done with the intention of finding meaning in positive and negative results. This meaning was intended to feed back into the cycle of policy and practice of the grand aims.

The aims relating to developing and delivering a methodology able to provide the type and quality of data demanded by the grand aims have all been achieved too. The methodology uses the 'gold standard' of the randomised control design to provide data of high reliability, which means that any positive findings have good potential to generalise. In addition the types of data gathered are varied between quantitative and qualitative which gives ecological validity and strength to the research design. A methodology of this type is seldom used for investigations into group effects in community settings and is a significant achievement in itself. The reflections made on the findings have been conducted in the context of relevant literature and been illuminated by key studies, and thoughtful contributions to theory and practice have been possible. These features mean that the findings are of sufficient quality to justify consideration for informing policy. Although the AGI intervention was not able to demonstrate greater gains than the CSG Intervention, it is possible to say that the AGI was at least as good as the CSG. It was more carefully constructed, took more account of the children members, was more onerous to deliver, but it was also more interesting, more enjoyable to be part of, more creative and more insightful. As far as the children are concerned, it may well be true that they felt no more benefit being part of the AGI group than the CSG group, but if the insights from a social constructivist perspective are to carry any weight, the children are only part of the solution. The teachers and the school environment are huge considerations in the phenomenon of disaffection, and leading AGI intervention is much more likely to change teacher practice than leading the CSG Intervention. So with both equal in terms of benefit to children, the success of the project has been to develop an intervention that is available to teachers, but will involve them in a change of outlook in order to deliver. If this change is towards a more therapeutic outlook in schools then the outcome is a good one

The Grand Aims (again):

Consideration of whether the AGI study has fulfilled its grand aims can only be done after consideration of whether it has fulfilled its other aims. For this reason the grand aims are re-stated and the pragmatic and research aims brought to bear on the success of the project.

The AGI study aimed to advance provision in youth disaffection, specifically: advances in the practice of service delivery in school, advances in techniques available to practitioners dealing with the challenge of disaffection, and advances in the behaviour profile of the children who are themselves disaffected. These advances are judged to have been made in a small way. Although meeting the criteria of scale in devising an intervention for disaffection, the practical application of the AGI interventions has not been commensurate with the scale of need, even in the North East of England. But this was not in the aims of the AGI study, and will perhaps forever remain something of an aspiration.

In addition, the aim was to provide advances in the theory of disaffection and advances in policy planning and development in youth disaffection in the school setting. In a small way, these aims have been judged to have been met. The theory-base has been enlarged through the reflections given to the research findings and these reflections tested by holding them up to key influential studies, dealing with the same issues. The links where theory impacts directly on practice have been explored and as a result the manner in which a programme of group intervention may happen in schools has been refined. In some cases, individual schools have modified their policy and practice to accommodate these insights from research and theory contained in the AGI study. Groupwork programmes are a continuing part of the AGI practice and a legacy of this research. The current state of practice is discussed below. The wider policy at area and national level is harder to influence, and what small effect the AGI study might have had on the national debate is hard to measure and anyway is miniscule. Nevertheless, the publication of the research findings (McArdle, Moseley, Quibell, Johnson, Allen, Hammal, & leCouteur 2002) partially fulfils the stated aim that research findings should be suitable for publication in a range of theoretical, professional and research journals.

The present state of the practice:

The Action GroupSkills Intervention (AGI) has been delivered in schools by The Learning Challenge (TLC) since 1994. TLC is a company limited by guarantee and a registered charity. The interventions evaluated by this research project were all delivered by one therapist, and up-scaling the interventions assumes that all therapists could be equally effective in delivering change to disaffected youth. Since the research took place, other therapists have been trained in the delivery of AGI sessions: in the academic year 2003 – 2004 TLC delivered groupwork programmes in six secondary schools in the North East region. These programmes involved the delivery of AGI groupwork sessions to disaffected young people by TLC staff. In every case where there was AGI delivery, a group of school staff were trained in the skills necessary to deliver the AGI programme. After two terms of training, the school staff begin a delivery programme to disaffected pupils under the supervision of TLC staff. A school team typically consists of 12 practitioners (teachers, learning mentors, school health professionals, classroom assistants). Working in a pair each person has one hour of delivery a week timetabled into their teaching. Six pairs each working with 8 children in a group mean 48 children each term receive the AGI provision. Each pair has new children at the start of every term, giving a yearly total of 144 children. Working across year 6, 7 and 8, the schools where this work is happening are reaching the at-risk population of the entire lower end of their schools. Evaluation of the training to school staff has been conducted by the King's Fund and found to enhance teacher's experience of their own practice as well as that of delivering in groups (Quibell 2002). Children involved in AGI sessions run by school staff report reduced levels of bullying and truancy, making it a reasonable assumption that the groupwork conducted by school staff is delivering some of the same benefits as that delivered in the AGI study.

Section 7.4: Limitations of present Study

7.4.1: Limitations of Design

In some analyses, the chances of finding statistical significance were compromised by the numbers. The numbers were affected by attrition (particularly in the case of the parental data) and by the slicing of the dataset to examine the group effects: as reported in Chapter 3, a sample of 120 children equates to 80% power. Halving that sample will lower the power while maintaining the desired effect size or reduce the effect size to maintain the 80% power level. In these cases the effect sizes have been used to

measure the magnitude of effect, but in order to increase the chances of statistical significance greater numbers are needed.

The design could be regarded as lacking a commitment to qualitative data, and some of the more sensitive responses to intervention could well have been more adequately gathered using interviewing techniques.

7.4.2: Limitations of Outcome

The study could not differentiate between intervention conditions. Quite some discussion has gone into this finding, and without wanting to duplicate what has already been said, it is worth mentioning as a limitation the frustrated prediction that the more intensive intervention would give the greater impact on the behaviour of children. The intervention conditions were intentionally similar, with both emphasising the positive power of the group situation, but further thought and work needs to go into differentiating the active ingredients of the group dynamic.

The data from the parents was more marginal in the support it gave to the hypotheses. Delivering a school-based intervention with the intention to produce behavioural benefits in the home is a tall order, but it is clear that when the home environment is the explicit focus of the intervention, gains can be produced. This aspect of the intervention would bear further exploration because of the importance and rarity of such findings. Kolvin found home benefits as a result of deeper parental involvement (Kolvin, Garside, Nicol, Macmillan, Wolstenholme, & Leitch 1981), and this could provide the way forward in enhancing the marginal gains found in this study.

The limitations found in the observational data were multiplied by the lack of proper validation of the measures used with the population. This was an oversight in the design exacerbated by urgency at the point where observational measures were being considered. The research team has not encountered observational protocols validated for universal use in the classroom. To compensate, the protocols used in any future study will be well advised to pilot the procedure to establish inter-rater and concurrent validity. The observation data were compromised as a result, particularly in the case of the whole class observations.

Attendance data was subject to confounding variables not anticipated in the design, most notably the decline in attendance over any transition period between schools, making any comparison between later and earlier time points unrepresentative. For this reason any future study would be well advised to identify some kind of group matched for symptoms to be tracked alongside the study sample. This is despite the ethical and practical difficulties in doing so.

7.4.3: Limitations Arising from Bias

The AGI Study has been designed as a quantitative study, a quasi-experiment, using measures designed to adhere as closely as possible to formal, some might say traditional, notions of control to increase generalisability. A quasi experiment contains most, if not all, of the elements of a formal experiment but has limitations of randomisation or control. The design itself is limited as outlined above, however the choice of methodology is itself a limitation when one considers the biases behind the very choice of a particular methodological shape. These choices constrict the data and determine to a large extent the analysis. The position of the researcher has been made clear throughout the previous chapters through introduction and reflective commentary as well as an explicit statement of values in Chapter 4. This engagement with the issues of reflexivity has been recognised as important in establishing the credibility of the study. For this reason attempts have been made in the text to allow transparency about the investment of the researcher and the issues that have led to the choices made. Because this has been a process that is intended to be woven into the text through the devices mentioned above, it is hoped that less detail is demanded here.

What receives less attention in the text is the deeper bias of sympathy of the researcher to his subject group, and the personal investment of the researcher in the AGI study – his desire for positive outcomes. This was raised as an issue at the beginning of Chapter Six, where the conventions for the reporting of negative findings were set out. The researcher has invested time and effort in the development of a groupwork programme for children, indeed has established an organisation to work in school to spread the practice. The best outcome for the researcher and his organisation would be unanimously positive findings. The result of this is a bias towards looking for explanations for negative findings that relate to methodological and extraneous factors rather than the inability of the intervention to produce positive findings. Positive findings

themselves are more likely to be attributed to the intervention and less space is given to looking for alternative explanations. The measures taken to account for this bias culminated in an effort within Chapter Six to balance the discussion by toning down interpretations to reflect a more cautious approach.

The researcher has been engaged in a process of learning during the conduct of the project. Although not a large-scale study in the sense of anything approaching a representative sample, the scale has been an enormous challenge to manage, as has the data entry, collation and analysis. The researcher has been rather proud of his arts background and the discipline of a quantitative study has been a struggle and an education. Over and above this, the researcher has been frustrated with the instruments used and their unresponsiveness to change, and regrets the lack of qualitative data that could illuminate the individual process of children. The experience of working with the children included in the study has been rather at odds with the impersonal cohort structure of the analysis, the relationships built up in the intervention deliveries has not been represented in the data gathered, and there is a question as to whether a rigorous qualitative methodology running alongside the quantitative would have provided a more complete picture of personal change.

Section 7.5: Implications for Future Research

As mentioned above, the whole question of what makes an intervention effective does require more research. Having established the efficacy of these group interventions, there are no data from this study as to what the 'active ingredients' are. Is it the adult attention? Or is it the supportive group culture, or the shared sense of purpose? These are questions of great importance for disaffection. There is no feeling here to echo the usual researcher's cry for more RCTs, but there is a need for more research into the group as a generic entity. That is to say the group as opposed to group therapy of one type or another.

The finding that the content of the group's activity is not as important as the fact of the group's existence would benefit from replication, and if found to be robust, then the implications fully explored. It has been noted that outcome literature can generally find evidence that pretty much everything works for something. Could research such as that

contained in these pages demonstrate effects of the fact of intervention over the content of intervention?

Many years ago, Jerome Frank wrote about 'non-specific factors' as agents of effectiveness in the therapeutic relationship (Frank 1961). The personal power of the therapist, the beliefs of the therapist, the perceived credibility of the knowledge-base, the ritual of the procedures, the place of safety where the therapy takes place. Could the same be true of the power of the group? What are the non-specific factors effective in the group? Rather than deepen the groove of outcome research in the hope that scientific justification can be found to favour one form of group theory over another, it might be that what bears closer examination is the nature of the relationship the group facilitator strikes up with a group. A recent article describes the therapeutic encounter... 'I find myself engaged in an engrossing struggle where emotions run high and the relationship is repeatedly challenged. These actions cannot be reduced to a series of technical procedures, nor can they be scientifically proved, just as you cannot scientifically prove the value of a friendship or a marriage' (Marziller 2004). This may well be right, and the 'active ingredients' of the group encounter may not be amenable to examination, but the definition of quality will begin with the identification of good practice, and there is great potential benefit to disaffected youth from a further delineation of these factors.

References

- Ablon, S. L. 1996, "The Therapeutic Action of Play", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 35, no. 4, pp. 545-547.
- Achenbach, T. M. 1991a, *Integrative guide for the 1991 CBCL/4-18, YSR and TRF profiles* Burlington VT, University of Vermont Department of Psychiatry.
- Achenbach, T. M. 1991b, *Manual for the Child Behavior Checklist / 4-18 and 1991 Profile* Burlington VT, University of Vermont Department of Psychiatry.
- Achenbach, T. M. 1991c, *Manual for the Teacher's Report Form and 1991 Profile* Burlington VT, University of Vermont Department of Psychiatry.
- Achenbach, T. M. 1991d, *Manual for the Youth Self Report and 1991 Profile* Burlington VT, University of Vermont Department of Psychiatry.
- Achenbach, T. M. 1993, *Empirically based taxonomy: How to use syndromes and profile types derived from the CBCL/4-18, TRF and YSR* Burlington VT, University of Vermont Department of Psychiatry.
- Achenbach, T. M. 1995, "Diagnosis, assessment, and comorbidity in psychosocial treatment research", *Journal of Abnormal Child Psychiatry*, vol. 23, pp. 45-64.
- Adalbjarnardottir, S. 1993, "Promoting children's social growth in the schools: An intervention study*1", *Journal of Applied Developmental Psychology*, vol. 14, no. 4, pp. 461-484.
- Adelman, H. S. & Taylor, L. 2000, "Shaping the future of mental health in schools", *Psychology in the Schools*, vol. 37, no. 1, pp. 49-60.
- Alvarez, H. K. & Ollendick, T. 2003, "Evidence-based treatment," in *Encyclopaedia of clinical child and paediatric psychology*, T. Ollendick & C. E. Schroeder, eds., New York, Kluwer Academic/Plenum Publishers.
- Azima, F. J. C. 1996, "Status of Adolescent Research," in *Group Therapy with Children and Adolescents*, P. Kymiss & D. A. e. Halpern, eds., American Psychiatric Press, Washington D.C., pp. 369-405.
- Barbour, R. 2001, "Checklists for improving rigour in qualitative research: A case of the tail wagging the dog?", *British Medical Journal*, vol. 322, pp. 1115-1117.
- Barker Lunn, J. C. 1967, "The development of scales to measure junior school children's attitudes", *British Journal of Educational Psychology*, vol. 39, pp. 64-71.
- Barkham, M., Shapiro, D. A., Hardy, G. E., & Rees, A. 1999, "Psychotherapy in two-plus-one sessions: outcomes of a randomized controlled trial of cognitive-behavioral and psychodynamic-Interpersonal therapy for subsyndromal depression", *Journal of Consulting & Clinical Psychology*, vol. 67, no. 2, pp. 201-211.

Barkley, R. A., Fischer, M., Smallish, L., & Fletcher, K. 2004, "Young Adult Follow-Up of Hyperactive Children: Antisocial Activities and Drug Use", *Journal of Child Psychology and Psychiatry*, vol. 45, no. 2, pp. 195-211.

Barrera, M., Biglan, A., Taylor, T. K., Gunn, B. K., Smolkowski, K., Black, C., Ary, D. V., & Fowler, R. C. 2002, "Early elementary school intervention to reduce conduct problems: a randomised Trial with hispanic and non-hispanic children", *Prevention Science*, vol. 3, no. 2, pp. 83-94.

Barrett, P., Healy-Farrell, L., & March, J. S. 2004, "Cognitive-Behavioral Family Treatment of Childhood Obsessive-Compulsive Disorder: A Controlled Trial", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 43, no. 1, pp. 46-62.

Bayliss, V. 2003. *Opening Minds: taking stock*. London, Royal Society for the encouragement of Arts, Manufactures and Commerce.

Beck, A. T., Ward, C. H., Mendelson, M., Mack, J., & Erbaugh, S. 1961. An inventory for measuring depression. *Archives of General Psychiatry* 4, 561-571.

Becker, H. S. 1982, "Culture: a sociological view", *Yale Review*, vol. 71, pp. 513-527.

Becker, L. A. Effect Size. 2000. <http://web.uccs.edu/lbecker/psy590/es.htm>. Accessed on 20.10.2004

Beckman, M. 1990. Collaborative Learning: Preparation for the Workplace and Democracy. *College Teaching* 38(4), 128-133.

Beelman, A., Pflingsten, U., & Losel, F. 1994, "Effects of training social competence in children: A meta-analysis of recent evaluation studies", *Journal of Clinical Child Psychology*, vol. 23, pp. 260-271.

Belmart, S., Anderson, B., Lee, S., & Utting, D. 2002, *Youth at Risk? A national survey of risk factors, protective factors and problem behaviour among young people in England Scotland and Wales* London, Communities that Care.

Bennett, B., Rolheiser, C., & Stevahn, L. 1991. *Cooperative Learning: where heart Meets Mind*. Toronto, Educational Connections.

Bentall, R. P. 2003, *Madness explained: psychosis and human nature* London, Allen Lane.

Berrueta-Clement, J. R., Shweinhart, L. J., Barnett, W. S., & Epstein, A. S. 1984, *Changed Lives: The Effects Of The Perry Preschool Program On Youths Through Age 19*. Ypsilanti, MI., High/Scope Press,

Bhaskar, R. 1989, *Reclaiming Reality: a critical introduction to contemporary philosophy* London, Verso.

Bion, W. R. *Experiences in Groups and other papers*. 1961. London, Tavistock Press.

Bird, H. R. 1996, "Epidemiology of childhood disorders in a cross-cultural context", *Journal of Child Psychology & Psychiatry*, vol. 37, pp. 35-49.

Bogdan, R. & Taylor, S. J. 1975, *Introduction to Qualitative Research Methods: a phenomenological approach to the social sciences* New York, Wiley.

Bouffard, T., Marcoux, M., Vezeau, C., & Bordeleau, L. 2003, "Changes in self-perceptions of competence and intrinsic motivation among elementary schoolchildren", *British Journal of Educational Psychology*, vol. 73, pp. 171-186.

Bracken, B. A. 1992, *Multidimensional Self Concept Scale* Austin, Texas, PRO-ED inc.

Bracken, B. A., Barona, A., Bauermeister, J. J., Juehn Howell, K., Poggioli, L., & Puente, A. 1990, "Multinational validation of the spanish Bracken basic concept scale for cross cultural assessments", *Journal of School Psychology*, vol. 28, no. 4, pp. 325-341.

Brint, S., Contreras, M. F., & Matthews, M. T. 2001, "Socialization messages in primary schools: an organizational analysis", *Sociology of Education*, vol. 74, no. 3, pp. 157-180.

British Psychological Society. Code of Conduct, Ethical Principles, and Guidelines. 2004.

British Sociological Association. Statement of Ethical Practice. 2004.

Bronson, M. B. 1994, "The usefulness of an observational measure of young children's social and mastery behaviors in early childhood classrooms", *Early Childhood Research Quarterly*, vol. 9, no. 1.

Brown, R. 2000, *Group Processes: dynamics within and between groups* Malden. MA., Blackwell.

Bryman, A. 2001, *Social Research Methods* Oxford, Oxford University Press.

Butler, S. & Wintram, C. 1991. *Feminist Groupwork*. London, Sage.

Camodeca, M., Goossens, F. A., Schuengel, C., & Terwogt, M. M. 2003, "Links between social information processing in middle childhood and involvement in bullying", *Aggressive Behaviour*, vol. 29, no. 2, pp. 116-127.

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*, vol. 65, pp. 684-698.

Canham, H. & Emanuel, L. 2000, "Tied together feelings' Group psychotherapy with latency children: the process of forming a cohesive group", *Journal of Child Psychotherapy*, vol. 26, no. 2, pp. 281-302.

Carver, R. P. 1978. The case against statistical significance. *Harvard Educational Review* 48, 378-399.

Carver, R. P. 1993. The case against statistical significance, revisited. *Journal of Experimental Education* 61, 287-292.

Casey, R. J. & Berman, J. S. 1985, "The outcome of psychotherapy with children", *Psychological Bulletin*, vol. 98, pp. 388-400.

Chambers, E. G. 1964, (2edn) *Statistical Calculation for Beginners* Cambridge, Cambridge University Press

Chapman, J. W. & Tunmer, W. E. 1997, "A longitudinal study of beginning reading achievement and reading self-concept", *British Journal of Educational Psychology*, vol. 67, pp. 279-291.

Chapman, J. W., Tunmer, W. E., & Prochnow, J. E. 2000, "Early reading related skills and performance, reading self-concept, and the development of academic self-concept: a longitudinal study", *Journal of Educational Psychology*, vol. 92, no. 4, pp. 703-708.

Chen, X., Chang, L., & He, Y. 2003, "The peer group as context: mediating and moderating effects on relations between academic achievement and social functioning in Chinese children", *Child Development*, vol. 74, no. 3, pp. 710-727.

Chickering, A. W. & Gamson, Z. F. e. Applying the Seven Principles for Good Practice in Undergraduate Education. New Directions for Teaching and Learning No 47 . 1991. San Francisco, JosseyBass.

Christie, D., Tolmie, A., Howe, C., Topping, K., Thurston, A., Jessiman, E., Livingston, K., & Donaldson, C. The Impact of Collaborative Group work in Primary Classrooms and the Effects of Class Composition in Urban and rural Schools. Paper presented at an Internal Teaching and Learning Research Programme Conference . 2001.

Cicchetti, D. & Toth, S. L. 1995, "A developmental psychopathology perspective on child abuse and neglect.", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 34, pp. 541-565.

Cicirelli, V. G. 1969, *The Impact of Headstart: An evaluation of the effects of Headstart on children's cognitive and affective development* Washington D.C., Westinghouse Learning Corporation, Ohio University.

Cicourel, A. V. 1982, "Interviews, Surveys and the Problem of Ecological Validity", *American Sociologist*, vol. 17, pp. 11-20.

Clark, C., Dyson, A., & Millward, A. 1998. *Theorising Special Education*. London, Routledge.

Clark, C., Prior, M., & Kinsella, G. 2002, "The relationship between executive function abilities, adaptive behaviour, and academic achievement in children with externalising behaviour problems", *Journal of Child Psychology & Psychiatry*, vol. 43, no. 6, pp. 785-796.

Clarke, M. & Oxman, A. D. 1999, *Cochrane reviewer's handbook 4.0* [updated July 1999].. Oxford, UK, Cochrane Collaboration.

Coe, R. 2000, *What is an effect size?* Draft manuscript, <http://www.cemcentre.org/eeg/default.asp>. Accessed on 20.10.2004

Cohen, J. 1988, *Statistical power for the behavioural sciences* Hillsdale, NJ., Erlbaum,

Conduct Problems Prevention Research Group 2002a, "The implementation of the Fast Track Program: an example of a large scale prevention science efficacy trial", *Journal of Abnormal Child Psychiatry*, vol. 30, no. 1, pp. 1-18.

Conduct Problems Prevention Research Group 2002b, "Evaluation of the first 3 years of the Fast Track prevention Trial with children at high risk for adolescent conduct problems", *Journal of Abnormal Child Psychiatry*, vol. 30, no. 1, pp. 19-35.

Corbin, J. M. 1998, "Alternative Interpretations: Valid or not?", *Theory and Psychology*, vol. 8, pp. 121-128.

Cowen, E. L., Hightower, A. D., Pedro-Carroll, J. L., Work, W. C., Wyman, P. A., & Haffey, W. G. 1996, *School-based Prevention for Children at Risk* Washington, American Psychological Association.

Coyle, G. L. 1937, *Studies in Group Behaviour* New York, Association Press.

Croll, P. 1986, *Systematic Classroom Observation* London, Falmer Press.

Cronbach, L. 1975, "Beyond the two disciplines of scientific psychology", *American Psychologist*, vol. 30, pp. 116-127.

Curry, J. F., Wells, K. C., Lochman, J. E., Craighead, W. E., & Nagy, P. D. 2003, "Cognitive-Behavioral Intervention for Depressed, Substance-Abusing Adolescents: Development and Pilot Testing", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 42, no. 6, pp. 656-665.

D'Amico, A. & Cardaci, M. Relations between perceived self-efficacy, self-esteem and school achievement. *Psychological Reports* 92[3], 745-754. 2003.

Dadds, M. R., Holland, D. E., Laurens, K. R., Mullins, M., Barrett, P. M., & Spence, S. H. 1999, "Early intervention and prevention of anxiety disorders in children: results at 2-year follow-up", *Journal of Consulting & Clinical Psychology*, vol. 67, no. 1, pp. 145-150.

Deikman, A. 1982, *The Observing Self: mysticism and psychotherapy* Boston, Beacon.

Denham, S. A., Blair, K. A., DeMulder, E., Levitas, J., Sawyer, K., Auerbach-Major, S., & Queenan, P. 2003, "Preschool Emotional Competence: Pathway to Social Competence?", *Child Development*, vol. 74, no. 1, pp. 238-256.

Devers, K. J. 1999, "How will we know 'good' qualitative research when we see it? Beginning the dialogue in health services research", *Health Services Research*, vol. 34, pp. 1153-1188.

Dewey, J. 1897, "My Pedagogic Creed", *The School Journal*, vol. LIV, no. 3, pp. 77-80.

Diamond, G. S., Reis, B. F., Diamond, G. M., Siqueland, L., & Isaacs, L. 2002, "Attachment-Based Family Therapy for Depressed Adolescents: A Treatment Development Study", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 41, no. 10, pp. 1190-1196.

Diener, E. & Crandall, R. 1978, *Ethics in Social and Behavioural Research* Chicago, University of Chicago Press.

Dies, K. G. 2000, "Adolescent Development and a Model of Group Psychotherapy: Effective Leadership in the New Millennium", *Journal of Child and Adolescent Group Therapy*, vol. 10, no. 2, pp. 97-111.

Downing, A. C. 2004, *RCOMPARE* Newcastle upon Tyne, UK, Dept Psychology, University of Newcastle.

DSM-IV 1994, *Diagnostic and statistical manual of mental disorders*, Washington D.C., American Psychiatric Association.

Duerr, M. 1993, *Early mental health initiative: Year-end evaluation report* Chico, CA., California Department of Mental Health.

Dunn, S. E., Lochman, J. E., & Colder, C. R. 1997, "Social problem-solving skills in boys with conduct and oppositional defiant disorders", *Aggressive Behaviour*, vol. 23, no. 6, pp. 457-469.

Durlak, J. A. 1995, *School-based prevention programs for children and adolescents* London, Sage.

Durlak, J. A. & Wells, A. 1994, "An evaluation of secondary mental health programmes for children and adolescents", *Paper presented at the first annual Kansas conference of child clinical psychology*.

Durlak, J. A., Wells, A., Cotton, J. K., & Johnson, S. 1995, "Analysis of selected methodological issues in child psychotherapy research", *Journal of Clinical Child Psychology*, vol. 24, pp. 141-148.

Durlak, J. A. & Wells, A. M. 1997, "Primary Prevention Mental Health Programmes For Children And Adolescents: A Meta- Analytic Review.", *American Journal of Community Psychology*, vol. 25, no. 2, pp. 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*, vol. 26, no. 5, pp. 775-802.

Edwards, A. D. & Westgate, D. P. G. 1987, *Investigating Classroom Talk* Lewes, Falmer Press.

Edwards, A. G., Russell, L. T., & Stott, N. C. H. 1998, "Signal versus noise in the evidence base for medicine: An alternative to hierarchies of evidence?", *Family Practice*, vol. 15, pp. 319-322.

Elliott, R., Fischer, C. T., & Rennie, D. L. 1999, "Evolving guidelines for publication of qualitative research studies in psychology and related fields", *British Journal of Clinical Psychology*, vol. 38, pp. 215-229.

Emler, N. 2001, *Self-esteem: the costs and causes of low self-worth*. York, Joseph Rowntree Foundation.

Emunah, R. 1994, *Acting for Real: drama therapy process, technique, and performance* Levittown, PA, Brunner/Mazel.

Eysenck, H. J. 1952, "The effects of psychotherapy: An evaluation", *Journal of Consulting Psychology*, vol. 16, pp. 319-324.

Eysenck, H. J. 1975, *Manual of the Eysenck personality questionnaire* London, Hodder and Stoughton.

Farrington, D. P. 1996, *Understanding and preventing youth crime. Social policy research*, London, Joseph Rowntree Foundation.

Feyerabend, P. 1975, *Against Method* London, Verso.

Flanders, N. A. 1970, *Analyzing Teaching Behavior* Reading, Mass., Addison Wesley.

Fonagy, P., Moran, G. S., & Target, M. 1993, "Aggression and the psychological self", *International Journal of Psycho-Analysis*, vol. 74, pp. 471-485.

Fonagy, P., Target, M., Cottrell, D., Phillips, J., & Kurtz, Z. 2002, *What Works for Whom? A Critical Review of Treatments for Children and Adolescents* New York, Guilford Press,

Ford, T., Goodman, R., & Meltzer, H. 2003, "The British Child and Adolescent Mental Health Survey 1999: The Prevalence of DSM-IV Disorders", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 42, no. 10, pp. 1203-1211.

Frank, J. 1961, *Persuasion and Healing* Baltimore MD., Johns Hopkins University Press.

Giddens, G. N. 1984, *The Constitution of Society* Cambridge, Polity Press.

Ginott, H. G. 1961, *Group Psychotherapy with Children: the theory and practice of play-therapy* New York, McGraw-Hill.

Ginsburg, G. S. & Drake, K. L. 2002, "School-Based Treatment for Anxious African-American Adolescents: A Controlled Pilot Study", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 41, no. 7, pp. 768-775.

Glasser, P., Sarri, R., & Vinter, R. E. 1974, *Individual Change Through Small Groups*. New York, Free Press.

Glassman, U. & Kates, L. 1990, *Group Work: a humanistic approach* Newbury Park, CA, Sage.

Goetschius, G. W. 1969, *Working with Community Groups* London, Routledge and Keegan Paul.

Goode, E. 1996, The Ethics of Deception in Social Research: a case study. *Qualitative Sociology* 19, 11-33. 1996.

Gottfredson, G. D., Jones, E. M., & Gore, T. W. 2002, "Implementation and evaluation of a cognitive-behavioural intervention to prevent problem behaviour in a disorganised school", *Prevention Science*, vol. 3, no. 1, pp. 43-56.

Gottman J.M., Katz, L. F., & Hooven, C. 1997, *Meta-emotion: How families communicate emotinally* Erlbaum, Mahwah, NJ.

Great Britain Home Office 2001, *Criminal Statistics England and Wales 2001: Statistics relating to criminal proceedings fro the year 2001* London, The Stationary Office.

Greenberg, M. T. 2004, "Current and future challenges in school-based prevention: the researcher perspective", *Prevention Science*, vol. 5, no. 1, pp. 5-13.

Greenwood, J. D. 1994, *Realism, Identity and Emotion* London, Sage.

Grizenko, N. 1997, "Outcome of multimodal day treatment for children with severe behavior problems: a five-year follow-up", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 36, no. 7, pp. 989-997.

Halperin, J. M. & McKay, K. E. 1998, "Psychological Testing for Child and Adolescent Psychiatrists: A Review of the Past 10 Years", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 37, no. 6, pp. 575-584.

Hansen, S., Meissler, K., & Ovens, R. 2000, "Kids Together: A Group Play Therapy Model for Children with ADHD Symptomology", *Journal of Child and Adolescent Group Therapy*, vol. 10, no. 4, pp. 191-211.

Hattie, J., Biggs, & Purdey 1996, "Effects of learning skills interventions on student learning: a meta-analysis", *Review of Educational Research*, vol. 66, pp. 99-136.

Hattie, J. & Flandaca, G. 2004, *Improving the quality of statistical studies in the Journal of Adult Development and Aging*. Greensboro, University of North Carolina.

Hayward, C., Varady, S., Albano, A. M., Thienemann, M., Henderson, L., & Schatzberg, A. 2000, "Cognitive-Behavioral Group Therapy for Social Phobia in Female Adolescents: Results of a Pilot Study", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 39, no. 6, pp. 721-726.

Hedges, L. & Olkin, I. 1985, *Statistical Methods for Meta-Analysis* London, Sage.

Henderlong, J. & Lepper, M. R. 2002, "Effects of praise on children's intrinsic motivation: a review and synthesis", *Psychological Bulletin*, vol. 128, no. 5, pp. 774-795.

Hightower, A. D., Cowen, E. L., Spinell, A. P., Lotyczewski, B. S., Guare, J. C., Rohrbeck, C. A., & Brown, L. P. 1987, "The Child rating scale: The development and

psychometric refinement of a socio-emotional self-rating scale for young schoolchildren", *School Psychology Review*, vol. 16, pp. 239-255.

Hightower, A. D., Work, W. C., Cowen, E. L., Lotyczewski, B. S., Spinell, A. P., Guare, J. C., & Rohrbeck, C. A. 1986, "The Teacher-Child rating Scale: A brief objective measure of elementary children's school problem behaviours", *School Psychology Review*, vol. 15, pp. 393-409.

Hoagwood, K. 2000, "Commentary: The Dose Effect in Children's Mental Health Services", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 39, no. 2, pp. 172-175.

Holdaway, S. 1982, *An inside job: a case study of covert research on the police*. in Bulmer, M. (ed.) *Social Research Ethics*. London, Macmillan.

Holmberg, L., Baum, M., & Adami, H. O. 1999, "On the scientific inference from clinical trials", *Journal of Evaluation in Clinical Practice*, vol. 5, pp. 157-162.

Hopkins W.G. 1997, *Effect Size Calculation* <http://sportsci.org/resource/stats>, accessed on 23/07/2002.

Hopkins, W. G. 2000, *What is a p Value?* <http://sportsci.org/resource/stats>, accessed on 23/07/2002

Hopkins, W. G. 2002, *Statistical vs clinical vs practical significance* <http://www.sportsci.org/jour/0103/statistical>, accessed on 28/09/04.

Howard, B. L. & Kendall, P. C. 1996, "Cognitive-behavioural family therapy for anxiety disordered children: a multiple baseline evaluation", *Cognitive Therapy and Research*, vol. 20, pp. 423-443.

Hundert, J., Boyle, M. H., Cunningham, C. E., Duku, E., Heale, J., McDonald, J., & Offord, D. R. 1999, "Helping children adjust - a Tri-ministry Study: II. Program Effects", *Journal of Child Psychology and Psychiatry*, vol. 40, no. 7, pp. 1061-1074.

ICD-10 1993, *The ICD-10 classification of mental and behavioural disorders. Diagnostic criteria for research*, Geneva, WHO Press.

Jackson, E. 2002, "Mental health in schools: what about the staff?", *Journal of Child Psychotherapy*, vol. 28, no. 2, pp. 129-146.

Jennings, S. 1980, *Playtherapy with Children*. London, Routledge.

Jennings, S. 1987, *Dramatherapy: theory and practice for teachers and clinicians* London, Routledge.

Johnson, D. L. & Walker, T. 1987, "Primary prevention of behaviour problems in Mexican-American children", *American Journal of Community Psychology*, vol. 15, pp. 375-385.

Johnson, D. W., Johnson, R. T., & Smith, K. A. 1991, *Cooperative Learning: Increasing College Faculty Instructional Productivity. ASHE-FRIC Higher Educational Report No.4* Washington D.C., School of Education and Human Development, George Washington University.

Johnson, L., Bruhn, R., Winek, J., Krepps, J., & Wiley, K. 1999, "The use of child-centered play therapy and filial therapy with Head Start families: a brief report", *Journal of Marital And Family Therapy*, vol. 25, no. 2, pp. 169-176.

Jones, R. A. 2003, "The construction of emotional and behavioural difficulties", *Educational Psychology in Practice*, vol. 19, no. 2, pp. 147-157.

Kaftarian, S., Robertson, E., Compton, W., Watts-Davis, B., & Volkow, N. 2004, "Blending prevention research and practice in schools: critical issues and suggestions", *Prevention Science*, vol. 5, no. 1, pp. 1-5.

Kagitcibasli, C., Sunar, D., & Bekman, S. 2001, "Long-term effects of early intervention: Turkish low-income mothers and children", *Journal of Applied Developmental Psychology*, vol. 22, no. 4, pp. 333-361.

Kataoka, S. H., Stein, B. D., Jaycox, L. H., Wong, M., Escudero, P., Tu, W., Zaragoza, C., & Fink, A. 2003, "A School-Based Mental Health Program for Traumatized Latino Immigrant Children", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 42, no. 3, pp. 311-318.

Kazdin, A. E. 2000, "Developing a research agenda for child and adolescent psychotherapy", *Archives of General Psychiatry*, vol. 57, pp. 829-836.

Kazdin, A. E., Bass, D., Ayers, W. A., & Rogers, A. 1990, "Empirical and clinical focus of child and adolescent psychotherapy research", *Journal of Consulting & Clinical Psychology*, vol. 58, pp. 729-740.

Keeves, J. P. 1988, *Educational Research, Methodology and Measurement: an International handbook*. Oxford, Pergammon Press.

Kolvin, I., Garside, R. F., Nicol, R., Macmillan, A., Wolstenholme, F., & Leitch, I. M. 1981, *Help Starts Here: The maladjusted child in the ordinary school* London, Tavistock Publications.

Konopka, G. 1963, *Social Group Work: a helping process* Englewood Cliffs, N.J., Prentice Hall.

Kreitler, S., Zigler, E., Kagan, S., Olsen, D., Weissler, K., & Kreitler, H. 1995, "Cognitive and motivational determinants of academic achievement and behaviour in third and fourth grade disadvantaged children", *British Journal of Educational Psychology*, vol. 65, pp. 297-316.

Kusche, C. A. & Greenberg, M. T. 1994, *The PATHS Curriculum* Seattle, WA, Developmental Research and Programs.

Lally, J. R., Mangione, P. L., Honig, A. S., & Wittner, D. S. 1988, "The Syracuse University Family Development Research Program Research Project: Long-range impact of an early intervention with low-income children and their families." in D. R. Powel, ed., *Annual advances in applied developmental psychology: Vol 3. Parent education in early childhood intervention: emerging directions in theory, research and practice.*, Norwood, NJ, Ablex. pp. 79-104.

Lane, K. L., Beebe-Frankenberger, M. E., Lambros, K. M., & Pierson, M. 2001, "Designing effective interventions for children at-risk for antisocial behavior: an integrated model of components necessary for making valid inferences", *Psychology in the Schools*, vol. 38, no. 4, pp. 365-379.

Lane, K. L., Umbreit, J., & Beebe-Frankenberger, M. E. 1999, "Functional assessment research on students with or at-risk for ADD; 1990 to the present", *Journal of Positive Behavior Interventions*, vol. 2, pp. 101-111.

Lasch, C. 1978, *The culture of narcissism: American life in an age of diminishing expectations* New York, Norton.

Laslett, R. 1983, *Changing perceptions of maladjusted children 1945 - 1981* Portishead, Association of Workers for Maladjusted Children.

Lazar, I. & Darlington, R. 1982, "Lasting Effects Of Early Education: A Report For The Consortium Of Longitudinal Studies. *Monographs Of Society For research in Child Development*, vol. 47, no. 1, pp. 1-51.

Leblanc, M. & Ritchie, M. 2001, "A meta-analysis of play therapy outcomes", *Counselling Psychology Quarterly*, vol. 14, no. 2, pp. 149-163.

Lengua, L. J. 2003, "Associations among emotionality, self-regulation, adjustment problems, and positive adjustment in middle childhood", *Journal of Applied Developmental Psychology*, vol. 24, no. 5, pp. 595-618.

Levi, D. 2001, *Group Dynamics for Teams* Thousand Oaks, Calif., Sage.

Levitt, E. E. 1957, "The results of psychotherapy with children: an evaluation", *Journal of Consulting & Clinical Psychology*, vol. 21, pp. 189-196.

Lewin, K. 1948, *Resolving Social Conflicts: selected papers on group dynamics* New York, Harper & Row.

Lewis, T. J., Powers, M. J. K., & Newcomer, L. L. 2002, "Reducing problem behaviors on the playground: an investigation of the application of schoolwide positive behavior supports", *Psychology in the Schools*, vol. 39, no. 2, pp. 181-190.

Lindsay, G., Dockrell, J., Letchford, B., & Mackie, C. 2003, "Self esteem of children with specific speech and language difficulties", *Child Language Teaching and Therapy*, vol. 20, pp. 125-143.

Lipsey, M. W. 1992, *Juvenile delinquency treatment: a meta-analytic inquiry into the variability of effects* in Cook, T. D., Cooper, H., Cordray, D. S., Hartmann, H., Hedges, L.

V., Light, R. J., Louis, T. A., and Mosteller, F. eds. *Meta-analysis for explanation* New York, Russell Sage Foundation.

Lochman, J. E. 1992, "Cognitive-behavioral intervention with aggressive boys: three-year follow-up and preventive effects", *Journal of Consulting & Clinical Psychology*, vol. 60, no. 3, pp. 426-432.

Lochman, J. E., Cole, J. D., Underwood, M. K., & Terry, R. 1993, "Effectiveness of a social relations intervention program for aggressive and nonaggressive, rejected children", *Journal of Consulting & Clinical Psychology*, vol. 61, no. 6, pp. 1053-1058.

Lomonaco, S., Scheidlinger, S., & Aronsen, E. 2000, "Five Decades of Children's Group Treatment - An Overview", *Journal of Child and Adolescent Group Therapy*, vol. 10, no. 2, pp. 77-96.

Lyle, S. 1996, "An analysis of collaborative group work in the primary school and the factors relevant to its success" *Language and Education* 10(1), 13-31.

Maddux, J.E. 2002, *Stopping the madness: positive psychology and the deconstruction of the illness ideology* In Snyder, C.R. & Lopez, S.J. *Handbook of Positive Psychology*, New York, OUP.

MacLennan, B. W. 2000, "The Future of Adolescent Psychotherapy Groups in the New Millennium", *Journal of Child and Adolescent Group Therapy*, vol. 10, no. 2, pp. 67-75.

Macnaughton, R. J. 1996, "Numbers, scales and qualitative research", *The Lancet*, vol. 347, pp. 1099-1100.

Mahler, M. S. 1971, "A study of separation-individuation process and its possible application to borderline phenomena in the psychoanalytic situation" *The Psychoanalytic Study of the Child* 26, 403-424. 1971.

Marsh, H., Parada, R., Seeshing-Yeung, A., & Healey, J. 2001, "Aggressive school troublemakers and victims: a longitudinal model examining the pivotal role of self-concept", *Journal of Educational Psychology*, vol. 93, no. 2, pp. 411-419.

Marsh, H. W. 1990, *Self-Description Questionnaire-II* San Antonio, TX., Psychological Corporation.

Marziller, J. 2004, "The myth of evidence-based psychotherapy", *The Psychologist*, vol. 17, no. 7, pp. 392-395.

Maslow, A. H. 1968, *Toward a psychology of being (2nd ed)* Princeton N.J., Van Nostrand.

Mays, N. & Pope, C. 1995, "Rigour and qualitative research", *British Medical Journal*, vol. 311, pp. 109-112.

McArdle, P., Moseley, D., Quibell, T., Johnson, R., Allen, A., Hammal, D., & leCouteur, A. 2002, "School-based indicated prevention: a randomised trial of group therapy", *Journal of Child Psychology & Psychiatry*, vol. 43, no. 6, pp. 705-712.

McIntosh, D. E., Rizza, M., & Bliss, L. 2000, "Implementing empirically supported interventions: teacher-child interaction therapy", *Psychology in the Schools*, vol. 37, no. 5, pp. 453-462.

McKey, R. H., Candelli, L., Ganson, H., Barre, H., McConbey, C., & Plantz, M. C. 1985, *The Impact Of Headstart On Children Families And Communities: Final Report of The Headstart Evaluation , Synthesis And Utilisation Project* Washington DC, Government Printing Office.

Meichenbaum, D. 1995, "The evolution of a cognitive behavior therapist", in J.K. Zeig (ed), *The Evolution of Psychotherapy: the third conference*, New York, Brunner/Mazel. pp. pp95-104.

Meller, P. J., Laboy, W., Rothwax, Y., Fritton, J., & Mangual, J. 1994, *Community School District Four: Primary Mental Health Project, 1990-1994* New York. Community School District #4.

Meltzer, H., Gatward, R., Goodman, R., & Ford, T. 2000, *Mental Health of Children and Adolescents in Great Britain*, London, The Stationary Office,

Mies, M. 1993, *Towards a Methodology for Feminist Research* Sage, London.

Moreno, J. L. 1964, *Psychodrama: Volume 1 (2nd edn)* New York, Beacon.

Morgan, G. A., Harmon, R. J., & Gliner, J. A. 2001, "Ethical Problems and Principles in Human Research", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 40, no. 10, pp. 1231-1233.

Mosley, J. & Sonnet, H. 2002, *101 Games For Self Esteem* Wisbech, Cambs, LDA.

Mosley, J. & Sonnet, H. 2003, *101 Games for Social Skills* Wisbech, Cambs, LDA.

Murray, T. R. 2000, *Comparing Theories of Child Development* Belmont, CA., Wadsworth Pub.

Myers, K. & Winters, N. C. 2002, "Ten-Year Review of Rating Scales. I: Overview of Scale Functioning, Psychometric Properties, and Selection", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 41, no. 2, pp. 114-122.

National Child Development Study. 2000, *BCS70-1970 British Cohort Study: 1999/2000 follow-up*, Centre for Longitudinal Studies, London, Institute of Education,

Norwich Union Healthcare. 2004. *Growing Pains*. Guardian Society. 6-10-2004.

Offord, D. R. 1987, "Prevention of Behavioural and Emotional Disorders in Children.", *Journal of Child Psychology and Psychiatry*, vol. 28, pp. 9-20.

Offord, D. R., Boyle, M. H., Szatmari, P., Rae-Grant, N. I., Links, P. S., Cadman, D. T., Byles, J. A., Crawford, J. W., Munroe Blum, H., Byrne, C., Thomas, H., & Woodward, C. A. 1987, "Ontario Child Health Study:II. Six month prevalence of disorder and rates of service utilization", *Archives of General Psychiatry*, vol. 44, pp. 832-836.

Ollendick, T. & King, N. J. 2004, *Empirically Supported Treatments for Children and Adolescents: Advances Towards Evidence-Based Practice* London, John Wiley & Sons.

Ollendick, T. H. & King, N. J. 1998, "Empirically supported treatments for children with phobic and anxiety disorders: current status, *Journal of Clinical Child Psychology*, vol. 27, no. 2, pp. 156-167.

Orobio de Castro, B., Bosch, J. D., Veerman, J. W., & Koops, W. 2003, "The Effects of Emotion Regulation, Attribution, and Delay IPrompts on Aggressive Boys' Social Problem Solving", *Cognitive Therapy and Research*, vol. 27, no. 2, pp. 153-166.

Patterson, G. R., Reid, J. B., & Dishion, T. J. 1992, *Antisocial boys: A social interactional approach* Eugene OR, Castalia Press.

Persons, J. B. & Silberschatz, G. 1998, "Are results of randomised controlled trials useful to psychotherapists?", *Journal of Consulting & Clinical Psychology*, vol. 66, pp. 126-135.

Pierrehumbert, B., Bader, M., Milijkovitch, R., Mazet, P., Amar, M., & Halfon, O. 2002, "Strategies of emotion regulation in adolescents and young adults with substance dependence or eating disorders", *Clinical Psychology and Psychotherapy*, vol. 9, no. 6, pp. 384-394.

Pollock, K. M. & Kymissis, P. 2001, "The Future of Adolescent Group Therapy. An Analysis of Historical Trends and Current Momentum", *Journal of Child and Adolescent Group Therapy*, vol. 11, no. 1, pp. 3-11.

Potter, J. 1996, *Representing Reality: discourse, rhetoric and social construction* London, Sage.

Prinz, R. J. 2002, "The Fast Track Project: a seminal intervention efficacy trial", *Journal of Abnormal Child Psychiatry*, vol. 30, no. 1, pp. 61-64.

Quibell, T. 2002, *Evaluation of a staff training programme in therapeutic groupwork*, London, The King's Fund.

Rabiner, D., Cole, J. D., & Conduct Problems Prevention Research Group 2000, "Early attention problems and children's reading achievement: a longitudinal investigation", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 39, no. 7, pp. 859-867.

Reid, K. E. 1981, *From Character Building to Social Treatment* Westport, Co., Greenwood Press.

Reid, R. & Maag, J. W. 1994, "How many fidgets in a pretty much: a critique of behavior rating scales for identifying students with ADHD", *Journal of School Psychology*, vol. 32, no. 4, pp. 339-354.

Rennie, D. L. 2000, "Grounded theory methodology as methodological hermeneutics", *Theory and Psychology*, vol. 10, pp. 481-502.

Reynolds, A. J. & Robertson, D. L. 2003, "School-Based Early Intervention and Later Child Maltreatment in the Chicago Longitudinal Study", *Child Development*, vol. 74, no. 1, pp. 3-26.

Roberts, R. 1998, "Prevalence of psychopathology among children and adolescents", *American Journal of Psychiatry*, vol. 155, pp. 715-725.

Rogers, C. R. 1970. *Encounter Groups* Harmondsworth, Penguin.

Rogerson, S. 1996, *Successful Groupwork*. London, Kogan Page.

Rosenthal, R. 1966, *Experimenter effects in behavioral research* New York, Appleton-Century-Crofts,

Rotheram-Borus, M. J. & Duan, N. 2003, "Next Generation of Preventive Interventions", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 42, no. 5, pp. 518-526.

Rutter, M. 1967, "A children's behaviour questionnaire for completion by teachers: preliminary findings", *Journal of Child Psychology & Psychiatry*, vol. 8, pp. 1-11.

Rutter, M., Cox, A., Tupling, C., Berger, M., & Yule, W. 1975, "Attainment and adjustment in two geographical areas: 1. The prevalence of psychiatric disorder", *British Journal of Psychiatry*, vol. 126, pp. 493-509.

Rutter, M., Giller, H., & Hagell, A. 1998, *Antisocial behaviour by young people* Cambridge, Cambridge University Press.

Sakoda, J. M., Cohen, B. H., & Beall, G. 1954, "Test of significance for a series of statistical tests", *Psychological Bulletin*, vol. 51, pp. 172-175.

Salmon, P. 2003, "Judging research quality", *The Psychologist*, vol. 16: 1, pp. 24-27.

Salovey, P. & Mayer, J. D. 1990, "Emotional Intelligence" *Imagination, Cognition and Personality* 9, 185-211.

Saltzman, W. R., Steinberg, A. M., Layne, C. M., Aisenberg, E., & Pynoos, R. S. 2001, "A Developmental Approach to School-Based Treatment of Adolescents Exposed to Trauma and Traumatic Loss", *Journal of Child and Adolescent Group Therapy*, vol. 11, no. 2/3, pp. 43-56.

Schulman, L. 1979 *Skills of Helping Individuals and Groups*. Itasca Illinois, Peacock.

Schutz, A. 1962, *Collected Papers I: the problem of social reality* The Hague, Martinus Nijhof.

Schwartz, W. & Zalba, S. R. 1971, *The Practice of Group Work* New York, Columbia University Press.

Schweinhart, L. J. 1999, "Evaluation of early childhood programs", *Child & Adolescent Psychiatric Clinics of North America*, vol. 8, pp. 395-407.

Schweinhart, L. J. & Barnes, H. V. 1993, *The High/Scope Perry Preschool Study Through Age 27*. New York, High/Scope Education Foundation.

Seitz, V., Rosenbaum, L. K., & Apfel, N. H. 1985, "Effects Of Family Support Intervention: A Ten Year Follow-up.", *Child Development*, vol. 56, pp. 376-391.

Shadish, W. R., Matt, G. E., Navarro, A. M., Siegle, G., Crits-Christoph, P., Hazelrigg, M. D., Jorm, A. F., Lyons, L. C., Nietzel, M. T., Prout, H. T., Robinson, L., Smith, M. L., Svartberg, M., & Weiss, B. 1997, "Evidence that therapy works in clinically representative conditions", *Journal of Consulting & Clinical Psychology*, vol. 65, no. 3, pp. 355-365.

Sharry, J. 2001, *Solution-Focused Groupwork* London, Sage.

Sherr, L. & Sterne, A. 1999, "Evaluation of a counselling intervention in primary schools", *Clinical Psychology and Psychotherapy*, vol. 6, no. 4, pp. 286-296.

Slade, E. 2002, "Effects of school-based mental health programs on mental health service use by adolescents at school and in the community", *Mental Health Services Record*, vol. 4, no. 3, pp. 151-166.

Slavin, R. E. 1980, "Cooperative Learning" *Review of Educational Research* 50(2), 315-342.

Slavin, R. E. & Madden, N. A. 1989. "What works for students at-risk? A research synthesis" *Educational Leadership* 46[4], pp 4-13.

Smith, D. J. 1995, "Youth crime and conduct disorders: Trends, patterns and causal explanations," in *Psychosocial disorders in young people: time trends and their causes*, M. Rutter & Smith D.J., eds., Chichester, UK, Wiley. pp. 389-489.

Snyder, K. V., Kymissis, P., & Kessler, K. 1999, "Anger Management for Adolescents: Efficacy of Brief Group Therapy", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 38, no. 11, pp. 1409-1416.

Social Research Association. 2004. Ethical Guidelines.

Sohn, D. 1998, "Statistic significance and replicability: why the former does not presage the latter", *Theory and Psychology*, vol. 8, pp. 291-311.

Sourander, A. & Piha, J. 1998, "Three-year follow-up of child psychiatric inpatient treatment", *European Child & Adolescent Psychiatry*, vol. 7, no. 3, pp. 153-162.

Stevens, R. 1996, "Trimodal theory as a model for interrelating perspectives in psychology," in *Issues for Social Psychology*, R. E. Sapsford, ed., Milton Keynes, Open University. pp. 77-84.

Stewart, G. L., Manz, C. C., & Sims, H. P. 1999, *Team work and group dynamics* New York, Wiley.

Stiles, W. B. 1993, "Quality Control in Qualitative Research", *Clinical Psychology Review*, vol. 13, pp. 593-618.

Stilwell, B. M., Glavin, M., Kopta, S. M., Padgett, R. J., & Holt, J. W. 1997, "Moralization of Attachment: A Fourth Domain of Conscience Functioning", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 36, no. 8, pp. 1140-1147.

Stormont, M. 2002, "Externalizing behavior problems in young children: contributing factors and early intervention", *Psychology in the Schools*, vol. 39, no. 2, pp. 127-138.

Strauss, A., Schatzman, L., Ehrich, D., Bucher, R., & Sabshin, M. 1973, "The Hospital and its Negotiated Order," in G. Salaman & K. E. Thompson (eds.), *People and Organizations*, London, Longman.

Strayhorn, J. M. 1988. *The Competent Child*. New York, The Guilford Press.

Strayhorn, J. M. 2002a, "Self-Control: Theory and Research", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 41, no. 1, pp. 7-16.

Strayhorn, J. M. 2002b, "Self-Control: Toward Systematic Training Programs", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 41, no. 1, pp. 17-27.

The Children Act. 2004, *In press*, London, Great Britain Home Office.

The Learning Challenge. 2004, *Annual Report*, available through <http://www.learningchallenge.org.uk>

Thompson, A. 1990, *A Guide to Ethical Practice in Psychotherapy* New York, John Wiley and Sons.

Toren, P., Wolmer, L., Rosental, B., Eldar, S., Koren, S., Lask, M., Weizman, R., & Laor, N. 2000, "Case Series: Brief Parent-Child Group Therapy for Childhood Anxiety Disorders Using a Manual-Based Cognitive-Behavioral Technique", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 39, no. 10, pp. 1309-1312.

Tremblay, R. E. & Craig, W. 1990, "Developmental Crime Prevention," in D. P. Farrington & M. Tonry, eds., *Building a safer society: Strategic approaches to crime prevention*, vol. 19 Chicago, University of Chicago Press. pp. 151-236.

Turpin, G., Barley, V., Beail, N., Scaife, J., Smith, J. A., & Slade, P. 1997, "Standards for research projects and theses involving qualitative methods", *Clinical Psychology Forum*, vol. 108, pp. 3-7.

Von Wright, G. H. 1971, *Explanation and Understanding* London, Routledge.

Walsh, D. 1972, "Sociology and the social world," in Filmer P. et al., eds., *New Directions in Sociological Theory*, London, Collier-Macmillan.

- Warren, S. L., Oppenheim, D., & Emde, R. N. 1996, "Can Emotions and Themes in Children's Play Predict Behavior Problems?" *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 35, no. 10, pp. 1331-1337.
- Webster-Stratton, C., Reid, M. J., & Hammond, M. 2001, "Preventing conduct problems, promoting social competence: A parent and teacher training partnership in head start", *Journal of Clinical Child Psychology*, vol. 30, no. 3.
- Weikart, D. P. & Schweinhart, L. J. 1992, "High Scope Preschool Programme outcomes.," in McCord J. & Tremblay R.E., eds., *Preventing Anti-Social Behaviour, Interventions from birth through adolescence*, New York, Guildford. pp. 67-86.
- Weiss, B., Catron, T., Harris, V., & Phung, T. M. 1999, "The effectiveness of traditional child psychotherapy", *Journal of Consulting & Clinical Psychology*, vol. 67, no. 1, pp. 82-94.
- Weiss, E. & Weisz, J. R. 1995, "Effectiveness of Psychotherapy", *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 34, no. 8, pp. 971-972.
- Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. 1987, "Effectiveness of psychotherapy with children and adolescents: a meta-analysis for clinicians", *Journal of Consulting & Clinical Psychology*, vol. 55, pp. 542-549.
- Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. 1995, "Effects of psychotherapy for children and adolescents revisited: a meta-analysis of treatment outcome studies.", *Psychological Bulletin*, vol. 117, pp. 450-468.
- Wentzel, K. & Caldwell, K. 1997, "Friendships, Peer Acceptance and group membership: Relations to academic achievement in middle school", *Child Development*, vol. 68, no. 6, pp. 1198-1209.
- West, M. 1996. *A Handbook of Work Group Psychology*. New York, John Wiley.
- Wood, A., Harrington, R., & Moore, A. 1996, "Controlled Trial Of A Brief Cognitive-Behavioural Intervention In Adolescent Patients With Depressive Disorders." *Journal of Child Psychology and Psychiatry*, vol. 37, pp. 737-746.
- Wragg, E. C. 1994, *An Introduction to Classroom Observation* London, Routledge.
- Yalom, I. D. 1975, *The Theory of Practice of Group Psychotherapy* New York, Basic Books.
- Yardley, L. 1997, *Material Discourses on Health and Illness* London, Routledge.
- Yoshikawa, H. 1994, "Prevention As Cumulative Protection: Effects Of Early Family Support And Education On Chronic Delinquency and Its Risks." *Psychological Bulletin*, vol. 115, no. 1, pp. 28-54.
- Youth Justice Board UK 2003, *Youth Survey 2003*, London, MORI.

Zsolnai, A. 2002, "Relationship between children's social competence, learning motivation and school achievement", *Educational Psychology*, vol. 22, no. 3, pp. 317-329.

Appendix 1: The Dataset: Abbreviated Version

ID	COHORT	BOHTYPE	DOB	DOI	CONDITO	SEX	AGEMONTH	AGEYRS	PRES.C1	C1.TTO	C1.TTS	C1.TOW	C1.INTS	C1.TOEX	C1.EXTS	PRES.T1	T1.TTO	T1.TTS	T1.TOW
1A.01	1.00	1	08-Nov-1985	15-Sep-1986	1.00	.00	130.23	10.88	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.02	1.00	1	11-Nov-1985	15-Sep-1986	.00	.00	130.14	10.84	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.03	1.00	1	18-Jul-1986	15-Sep-1986	1.00	1.00	122.02	10.17	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.04	1.00	1	30-Jul-1986	15-Sep-1986	.00	1.00	121.56	10.13	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.05	1.00	1	16-Sep-1986	15-Sep-1986	1.00	1.00	121.91	10.08	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.06	1.00	1	12-Nov-1985	15-Sep-1986	.00	1.00	130.10	10.84	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.07	1.00	1	07-Feb-1986	15-Sep-1986	.00	1.00	127.34	10.60	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.08	1.00	1	20-Apr-1986	15-Sep-1986	1.00	.00	124.88	10.41	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.09	1.00	1	03-Aug-1986	15-Sep-1986	1.00	.00	121.43	10.12	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.10	1.00	1	20-Sep-1985	15-Sep-1986	.00	1.00	121.24	10.08	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.11	1.00	1	06-May-1986	15-Sep-1986	1.00	.00	124.29	10.36	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.12	1.00	1	#NULL!	15-Sep-1986	1.00	1.00	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.13	1.00	1	26-Apr-1986	15-Sep-1986	1.00	1.00	124.02	10.30	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.14	1.00	1	23-Mar-1986	15-Sep-1986	1.00	1.00	125.80	10.48	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.15	1.00	1	14-Jun-1985	15-Sep-1986	.00	.00	140.02	11.67	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.16	1.00	1	15-Jul-1986	15-Sep-1986	.00	.00	122.05	10.17	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1A.17	1.00	1	25-Jul-1986	15-Sep-1986	.00	1.00	121.72	10.14	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.01	1.00	2	27-Feb-1985	15-Sep-1986	.00	1.00	136.56	11.55	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.02	1.00	2	19-Jun-1985	15-Sep-1986	1.00	.00	134.99	11.34	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.03	1.00	2	16-Dec-1984	15-Sep-1986	1.00	.00	140.99	11.75	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.04	1.00	2	09-May-1985	15-Sep-1986	1.00	1.00	136.44	11.37	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.05	1.00	2	16-Oct-1984	15-Sep-1986	1.00	.00	142.86	11.82	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.06	1.00	2	08-Jun-1985	15-Sep-1986	1.00	1.00	140.22	11.69	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.07	1.00	2	27-Feb-1986	15-Sep-1986	1.00	1.00	126.59	10.55	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.08	1.00	2	22-Nov-1984	15-Sep-1986	.00	.00	141.77	11.81	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.09	1.00	2	#NULL!	15-Sep-1986	.00	1.00	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.10	1.00	2	06-Nov-1984	15-Sep-1986	.00	1.00	142.23	11.85	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.11	1.00	2	19-Jul-1985	15-Sep-1986	.00	1.00	134.21	11.18	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.12	1.00	2	06-Jul-1985	15-Sep-1986	1.00	.00	134.34	11.20	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.13	1.00	2	01-May-1985	15-Sep-1986	.00	1.00	138.51	11.54	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.14	1.00	2	10-Mar-1985	15-Sep-1986	.00	.00	136.22	11.52	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.15	1.00	2	#NULL!	15-Sep-1986	.00	1.00	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
1B.16	1.00	2	27-Apr-1985	15-Sep-1986	.00	1.00	137.66	11.47	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!
2A.01	2.00	2	20-Dec-1984	15-Jan-1987	1.00	.00	144.85	12.07	1.00	1.00	40.00	3.00	46.00	.00	32.00	1.00	.00	22.00	.00
2A.02	2.00	2	26-May-1984	15-Jan-1987	1.00	.00	136.73	11.84	1.00	14.00	46.00	6.00	51.00	7.00	50.00	1.00	32.00	59.00	3.00
2A.03	2.00	2	16-Sep-1984	15-Jan-1987	.00	1.00	148.04	12.33	1.00	36.00	59.00	8.00	57.00	10.00	62.00	1.00	4.00	40.00	.00
2A.04	2.00	2	14-Sep-1984	15-Jan-1987	.00	.00	148.04	12.34	1.00	17.00	46.00	4.00	46.00	7.00	50.00	1.00	7.00	48.00	.00
2A.05	2.00	2	27-Jul-1985	15-Jan-1987	.00	.00	137.86	11.47	1.00	13.00	46.00	1.00	39.00	4.00	44.00	1.00	23.00	55.00	9.00
2A.06	2.00	2	30-Apr-1986	15-Jan-1987	1.00	.00	140.56	11.71	1.00	21.00	51.00	3.00	51.00	3.00	46.00	1.00	12.00	51.00	2.00
2A.07	2.00	2	11-Dec-1984	15-Jan-1987	1.00	1.00	146.18	12.10	1.00	44.00	63.00	10.00	60.00	26.00	66.00	1.00	6.00	43.00	.00
2A.08	2.00	2	05-Oct-1984	15-Jan-1987	.00	.00	147.26	12.26	1.00	19.00	49.00	6.00	52.00	6.00	46.00	1.00	12.00	51.00	7.00
2A.09	2.00	2	#NULL!	15-Jan-1987	1.00	1.00	#NULL!	#NULL!	1.00	11.00	42.00	2.00	43.00	4.00	43.00	1.00	11.00	47.00	10.00
2A.10	2.00	2	31-Oct-1984	15-Jan-1987	.00	.00	146.50	12.21	1.00	16.00	47.00	6.00	52.00	6.00	46.00	1.00	.00	32.00	.00
2A.11	2.00	2	26-Feb-1986	15-Jan-1987	1.00	1.00	142.56	11.88	1.00	21.00	49.00	6.00	53.00	7.00	47.00	1.00	9.00	46.00	7.00
2A.12	2.00	2	16-Aug-1985	15-Jan-1987	1.00	1.00	136.94	11.41	1.00	6.00	37.00	2.00	43.00	2.00	36.00	1.00	6.00	42.00	2.00
2A.13	2.00	2	16-Aug-1985	15-Jan-1987	1.00	1.00	136.94	11.41	1.00	86.00	76.00	13.00	84.00	46.00	63.00	1.00	52.00	62.00	1.00
2A.14	2.00	2	26-May-1985	15-Jan-1987	1.00	.00	136.78	11.84	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	31.00	55.00	6.00
2A.15	2.00	2	04-Dec-1984	15-Jan-1987	1.00	1.00	146.38	12.11	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	50.00	61.00	7.00
2A.16	2.00	2	26-Feb-1986	15-Jan-1987	.00	.00	142.62	11.89	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	24.00	56.00	6.00
2B.01	2.00	1	04-Nov-1986	15-Jan-1987	.00	.00	122.38	10.20	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	17.00	54.00	13.00
2B.02	2.00	1	14-Aug-1986	15-Jan-1987	.00	.00	126.08	10.42	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	21.00	55.00	14.00
2B.03	2.00	1	21-Jan-1986	15-Jan-1987	1.00	1.00	131.81	10.96	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	13.00	48.00	3.00
2B.04	2.00	1	10-Dec-1985	15-Jan-1987	1.00	1.00	133.19	11.10	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	52.00	62.00	14.00
2B.05	2.00	1	19-Jan-1986	15-Jan-1987	.00	1.00	131.91	10.99	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	46.00	60.00	13.00
2B.06	2.00	1	16-Aug-1986	15-Jan-1987	1.00	1.00	125.21	10.43	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	64.00	74.00	13.00
2B.07	2.00	1	26-Jun-1986	15-Jan-1987	.00	1.00	126.62	10.55	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	.00	31.00	.00
2B.08	2.00	1	09-Dec-1986	15-Jan-1987	.00	.00	121.23	10.19	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	17.00	54.00	15.00
2B.09	2.00	1	25-Jun-1986	15-Jan-1987	.00	.00	126.72	10.58	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	50.00	66.00	15.00
2B.10	2.00	1	01-May-1986	15-Jan-1987	1.00	.00	126.53	10.71	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	41.00	63.00	31.00
2B.11	2.00	1	#NULL!	15-Jan-1987	1.00	.00	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	60.00	67.00	10.00
2B.12	2.00	1	26-Jun-1986	15-Jan-1987	1.00	1.00	131.84	10.96	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	107.00	77.00	15.00
2B.13	2.00	1	12-Nov-1986	15-Jan-1987	1.00	1.00	134.11	11.16	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	23.00	53.00	17.00
2B.14	2.00	1	05-Sep-1986	15-Jan-1987	1.00	1.00	124.36	10.30	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	37.00	58.00	4.00
2B.15	2.00	1	26-Sep-1986	15-Jan-1987	.00	1.00	135.89	11.30	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	40.00	69.00	16.00
2B.16	2.00	1	#NULL!	15-Jan-1987	.00	.00	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	76.00	69.00	31.00
2C.01	2.00	3	06-Apr-1985	15-Jan-1987	.00	1.00	141.44	11.79	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	63.00	71.00	4.00
2C.02	2.00	3	13-Dec-1984	15-Jan-1987	.00	.00	145.08	12.00	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	53.00	65.00	23.00
2C.03	2.00	3	31-May-1984	15-Jan-1987	.00	.00	139.53	11.63	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	16.00	54.00	12.00
2C.04	2.00	3	04-May-1985	15-Jan-1987	.00	.00	140.42	11.79	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	66.00	74.00	10.00
2C.05	2.00	3	#NULL!	15-Jan-1987	.00	1.00	#NULL!	#NULL!	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	19.00	51.00	6.00
2C.06	2.00	3	16-May-1985	15-Jan-1987	.00	1.00	141.96	11.83	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	13.00	46.00	2.00
2C.07	2.00	3	19-May-1985	15-Jan-1987	.00	1.00	139.93	11.66	.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	19.		

[illegible]

	M1.AGTS	M1.FATO	M1.FALMS	M1.FATS	M1.FHTO	M1.FHMS	M1.FHTS	M1.T.TO	M1.TMS	M1.TTS	PRES.C2	C2.TTO	C2.TTH	C2.TOIN	C2.INTS	C2.TOEX	C2.EXTS	PRES.T2	T2.TTO	T2.TTS
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	36.00	80.00	6.00	51.00	12.00	67.00	1.00	41.00	62.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	24.00	62.00	6.00	62.00	9.00	63.00	1.00	4.00	44.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	43.00	62.00	14.00	68.00	16.00	36.00	1.00	8.00	43.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	7.00	36.00	1.00	40.00	2.00	38.00	1.00	36.00	67.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	7.00	36.00	3.00	46.00	1.00	36.00	1.00	27.00	64.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	20.00	49.00	6.00	63.00	2.00	38.00	1.00	41.00	60.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	16.00	46.00	2.00	43.00	11.00	65.00	1.00	4.00	40.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	56.00	66.00	24.00	72.00	11.00	66.00	1.00	67.00	67.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	12.00	44.00	3.00	46.00	3.00	42.00	1.00	6.00	46.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	26.00	63.00	7.00	66.00	10.00	62.00	1.00	66.00	62.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	26.00	54.00	6.00	61.00	14.00	60.00	1.00	16.00	63.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	2.00	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	44.00	60.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	2.00	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	36.00	67.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	2.00	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	86.00	72.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	2.00	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	41.00	65.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	2.00	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	43.00	64.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	2.00	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	31.00	66.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	6.00	36.00	.00	34.00	2.00	38.00	1.00	.00	31.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	33.00	66.00	3.00	46.00	16.00	60.00	1.00	16.00	64.00
#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	#NULJ1	1.00	28.00	55.00	10.00	64.00	6.00	62			

T2.TOWH	T2.INT3	T2.TOE1	T2.EXT3	PRES.Y2	Y2.TT0	Y2.TT3	Y2.TOWH	Y2.INT3	Y2.TOE1	Y2.EXT3	PRES.M2	M2.S0.T0	M2.S0.M3	M2.S0.T3	M2.CO.T0	M2.CO.M3	M2.CO.T3	M2.AF.T3	M2.AC.T0	M2.AC.M3
8.00	54.00	12.00	82.00	1.00	47.00	85.00	16.00	87.00	83.00	55.00	1.00	57.00	77.00	65.00	60.00	77.00	65.00	85.00	88.00	80.00
1.00	43.00	2.00	82.00	1.00	84.00	86.00	16.00	87.00	82.00	54.00	1.00	78.00	95.00	63.00	71.00	94.00	84.00	84.00	89.00	91.00
1.00	43.00	3.00	80.00	1.00	85.00	86.00	16.00	89.00	83.00	83.00	1.00	80.00	80.00	60.00	85.00	115.00	40.00	38.00	82.00	111.00
3.00	48.00	4.00	81.00	1.00	84.00	73.00	26.00	86.00	27.00	68.00	1.00	54.00	74.00	87.00	82.00	66.00	73.00	82.00	80.00	81.00
16.00	86.00	.00	38.00	4.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	4.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL
2.00	46.00	16.00	80.00	1.00	82.00	82.00	23.00	86.00	14.00	84.00	1.00	84.00	84.00	61.00	69.00	90.00	57.00	54.00	71.00	96.00
4.00	81.00	.00	38.00	1.00	71.00	86.00	21.00	84.00	26.00	86.00	1.00	76.00	100.00	60.00	75.00	100.00	50.00	53.00	81.00	130.00
17.00	86.00	16.00	86.00	1.00	86.00	82.00	22.00	86.00	16.00	36.00	1.00	53.00	73.00	86.00	85.00	84.00	81.00	81.00	80.00	84.00
3.00	49.00	16.00	82.00	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL
7.00	86.00	26.00	86.00	1.00	84.00	83.00	11.00	83.00	39.00	79.00	1.00	77.00	102.00	49.00	84.00	83.00	81.00	84.00	80.00	81.00
.00	37.00	10.00	80.00	1.00	29.00	46.00	1.00	32.00	16.00	86.00	1.00	84.00	112.00	42.00	78.00	106.00	47.00	43.00	95.00	125.00
29.00	73.00	.00	39.00	1.00	46.00	88.00	14.00	86.00	83.00	53.00	1.00	71.00	83.00	55.00	72.00	96.00	53.00	52.00	84.00	85.00
1.00	43.00	26.00	85.00	1.00	42.00	83.00	13.00	85.00	83.00	83.00	1.00	74.00	87.00	82.00	54.00	88.00	71.00	47.00	88.00	91.00
3.00	48.00	53.00	84.00	1.00	36.00	80.00	5.00	43.00	88.00	39.00	1.00	70.00	81.00	58.00	67.00	87.00	98.00	52.00	85.00	96.00
19.00	67.00	4.00	86.00	1.00	44.00	83.00	11.00	49.00	16.00	82.00	1.00	59.00	78.00	84.00	64.00	86.00	71.00	83.00	55.00	75.00
6.00	54.00	12.00	82.00	1.00	36.00	49.00	12.00	61.00	83.00	56.00	1.00	85.00	116.00	40.00	83.00	112.00	42.00	34.00	84.00	114.00
3.00	48.00	19.00	81.00	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL
.00	36.00	.00	39.00	1.00	46.00	85.00	12.00	54.00	12.00	52.00	1.00	54.00	74.00	87.00	82.00	80.00	63.00	59.00	73.00	100.00
.00	36.00	4.00	56.00	1.00	84.00	82.00	24.00	82.00	9.00	48.00	1.00	72.00	84.00	54.00	67.00	87.00	88.00	82.00	73.00	100.00
.00	36.00	.00	42.00	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL
10.00	80.00	34.00	85.00	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	1.00	88.00	118.00	36.00	87.00	114.00	38.00	51.00	78.00	107.00
7.00	86.00	3.00	54.00	1.00	40.00	82.00	11.00	49.00	11.00	82.00	1.00	57.00	76.00	66.00	88.00	96.00	61.00	80.00	81.00	81.00
10.00	80.00	.00	40.00	1.00	39.00	82.00	11.00	53.00	13.00	81.00	1.00	81.00	106.00	45.00	73.00	97.00	52.00	45.00	89.00	93.00
3.00	81.00	4.00	52.00	1.00	81.00	89.00	30.00	72.00	24.00	85.00	1.00	86.00	116.00	40.00	89.00	96.00	59.00	48.00	96.00	116.00
8.00	86.00	4.00	56.00	1.00	26.00	46.00	5.00	41.00	7.00	46.00	1.00	73.00	96.00	63.00	89.00	90.00	57.00	51.00	71.00	96.00
16.00	86.00	28.00	87.00	1.00	34.00	49.00	2.00	35.00	21.00	82.00	1.00	80.00	122.00	35.00	73.00	97.00	52.00	36.00	56.00	76.00
.00	39.00	12.00	58.00	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	1.00	79.00	115.00	40.00	81.00	108.00	44.00	44.00	70.00	85.00
.00	36.00	9.00	96.00	1.00	24.00	44.00	11.00	83.00	8.00	43.00	1.00	86.00	88.00	58.00	85.00	84.00	61.00	59.00	88.00	91.00
3.00	81.00	.00	42.00	1.00	111.00	76.00	39.00	76.00	28.00	72.00	1.00	84.00	84.00	61.00	81.00	78.00	65.00	63.00	49.00	88.00
2.00	48.00	16.00	80.00	1.00	83.00	83.00	19.00	82.00	18.00	39.00	1.00	87.00	117.00	39.00	71.00	94.00	54.00	54.00	73.00	98.00
1.00	45.00	22.00	87.00	1.00	83.00	81.00	16.00	82.00	19.00	63.00	1.00	91.00	124.00	34.00	86.00	98.00	58.00	48.00	87.00	90.00
8.00	87.00	.00	49.00	8.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL
15.00	84.00	27.00	86.00	1.00	84.00	81.00	13.00	85.00	30.00	72.00	1.00	85.00	85.00	80.00	86.00	86.00	58.00	59.00	80.00	81.00
2.00	49.00	.00	42.00	1.00	30.00	47.00	16.00	85.00	3.00	38.00	1.00	72.00	84.00	54.00	78.00	102.00	49.00	47.00	74.00	102.00
2.00	48.00	16.00	86.00	1.00	39.00	81.00	12.00	81.00	81.00	82.00	1.00	87.00	117.00	39.00	71.00	94.00	54.00	45.00	82.00	111.00
.00	38.00	9.00	56.00	1.00	45.00	85.00	12.00	54.00	13.00	53.00	1.00	86.00	115.00	40.00	73.00	97.00	52.00	47.00	83.00	84.00
.00	39.00	.00	42.00	1.00	28.00	46.00	9.00	47.00	8.00	47.00	1.00	86.00	115.00	40.00	81.00	108.00	44.00	38.00	90.00	120.00
1.00	45.00	.00	42.00	1.00	54.00	85.00	25.00	83.00	9.00	48.00	1.00	74.00	87.00	52.00	86.00	117.00	39.00	52.00	78.00	105.00
.00	36.00	.00	42.00	1.00	21.00	42.00	10.00	48.00	1.00	32.00	1.00	81.00	108.00	45.00	75.00	100.00	50.00	51.00	73.00	100.00
3.00	81.00	17.00	81.00	1.00	78.00	86.00	24.00	87.00	23.00	84.00	1.00	72.00	84.00	54.00	67.00	87.00	88.00	61.00	86.00	84.00
.00	36.00	.00	42.00	1.00	32.00	46.00	10.00	46.00	7.00	46.00	1.00	85.00	114.00	41.00	87.00	118.00	38.00	45.00	84.00	114.00
8NULL	8NULL	8NULL	8NULL	1.00	13.00	37.00	3.00	38.00	7.00	46.00	1.00	90.00	122.00	36.00	82.00	126.00	33.00	27.00	94.00	127.00
.00	35.00	.00	42.00	1.00	28.00	46.00	13.00	82.00	9.00	49.00	1.00	89.00	120.00	37.00	82.00	111.00	43.00	40.00	92.00	123.00
10.00	80.00	.00	40.00	1.00	36.00	80.00	9.00	80.00	11.00	80.00	1.00	70.00	91.00	58.00	81.00	109.00	44.00	45.00	81.00	110.00
.00	36.00	.00	40.00	1.00	54.00	80.00	23.00	86.00	14.00	54.00	1.00	81.00	81.00	83.00	70.00	92.00	56.00	56.00	77.00	106.00
7.00	86.00	48.00	76.00	1.00	34.00	49.00	7.00	46.00	12.00	82.00	2.00	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL	8NULL
4.00	83.00	2.00	82.00	1.00	26.00	46.00	12.00	81.00	4.00	49.00	1.00	88.00	120.00	37.00	77.00	103.00	46.00	34.00	82.00	123.00
22.00	86.00	18.00	82.00	1.00	70.00	85.00	21.00	84.00	14.00	84.00	1.00	80.00	80.00	63.00	70.00	92.00	55.00	56.00	75.00	104.00
4.00	83.00	.00	42.00	1.00	114.00	79.00	34.00	71.00	33.00	77.00	1.00	83.00	83.00	61.00	85.00	84.00	61.00	63.00	86.00	86.00
10.00	80.00	2.00	82.00	1.00	38.00	81.00	19.00	86.00	7.00	48.00	1.00	84.00	84.00	61.00	79.00	108.00	46.00	44.00	75.00	104.00
12.00	82.00	4.00	86.00	1.00	27.00	46.00	10.00	46.00	12.00	54.00	1.00	76.00	86.00	61.00	70.00	92.00	55.00	51.00	74.00	102.00
6.00	54.00	19.00	87.00	1.00	27.00	46.00	9.00	80.00	7.00	46.00	1.00	77.00	102.00	49.00	72.00	98.00	53.00	49.00	71.00	86.00
13.00	83.00	21.00	82.00	1.00	43.00	54.00	15.00	87.00	14.00	54.00	1.00	88.00	120.00	37.00	76.00	102.00	49.00	36.00	56.00	116.00
7.00	85.00	29.00	82.00	1.00	36.00	80.00	12.00	84.00	14.00	54.00	1.00	86.00	120.00	37.00	84.00	114.00	41.00	43.00	74.00	102.00
12.00	82.00	47.00	88.00	1.00	37.00	81.00	10.00	81.00	7.00	46.00	1.00	76.00	100.00	50.00	67.00	87.00	59.00	54.00	85.00	86.00
4.00	81.00	.00	39.00	1.00	6.00	30.00	3.00	38.00	2.00	33.00	1.00	79.00	105.00	47.00	75.00	100.00	50.00	52.00	71.00	86.00
6.00	54.00	.00	42.00	1.00	28.00	46.00	8.00	46.00	5.00	42.00	1.00	86.00	115.00	40.00	86.00	117.00	39.00	40.00	87.00	117.00
17.00	86.00	4.00	56.00	1.00	6.00	36.00	2.00	36.00	2.00	36.00	1.00	72.00	84.00	54.00	67.00	87.00	88.00	84.00	74.00	102.00
28.00	72.00	.00	42.00	1.00	45.00	54.00	16.00	57.00	5.00	42.00	1.00	74.00	87.00	52.00	67.00	87.00	89.00	61.00	86.00	83.00
13.00	83.00	48.00	83.00	1.00																

M2.AC.TS	M2.FA.TO	M2.FA.HS	M2.FA.TS	M2.PH.TO	M2.PH.HS	M2.PH.TS	M2.T.TO	M2.T.HS	M2.T.TS	PREB.CS	C3.TTO	C3.TTS	C3.TOW	C3.NTS	C3.OEX	C3.EXTS	PREB.TS	T3.TTO	T3.TTS	T3.TOW
63.00	71.00	65.00	68.00	80.00	82.00	82.00	300.00	78.00	87.00	1.00	38.00	80.00	8.00	67.00	7.00	50.00	1.00	32.00	88.00	7.00
56.00	73.00	67.00	69.00	73.00	83.00	82.00	415.00	89.00	87.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	16.00	53.00	9.00
43.00	80.00	105.00	47.00	88.00	80.00	85.00	476.00	105.00	47.00	1.00	24.00	51.00	1.00	51.00	10.00	52.00	1.00	13.00	48.00	3.00
63.00	68.00	82.00	62.00	85.00	88.00	87.00	358.00	74.00	67.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	28.00	54.00	5.00
#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	13.00	44.00	6.00	53.00	3.00	41.00	1.00	34.00	57.00	17.00
53.00	84.00	87.00	52.00	70.00	100.00	80.00	431.00	83.00	55.00	1.00	16.00	45.00	1.00	51.00	3.00	41.00	1.00	17.00	50.00	8.00
43.00	77.00	81.00	56.00	76.00	107.00	45.00	458.00	101.00	49.00	1.00	25.00	52.00	7.00	56.00	13.00	35.00	1.00	6.00	43.00	4.00
54.00	81.00	85.00	53.00	78.00	95.00	83.00	385.00	84.00	81.00	1.00	64.00	70.00	21.00	71.00	16.00	64.00	1.00	56.00	66.00	22.00
#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	7.00	38.00	1.00	39.00	2.00	40.00	1.00	12.00	51.00	8.00
83.00	87.00	89.00	51.00	63.00	86.00	59.00	405.00	85.00	59.00	1.00	18.00	48.00	1.00	51.00	4.00	48.00	1.00	21.00	52.00	3.00
33.00	87.00	113.00	41.00	81.00	116.00	43.00	517.00	116.00	39.00	1.00	15.00	46.00	2.00	43.00	6.00	48.00	1.00	8.00	48.00	4.00
80.00	83.00	105.00	47.00	80.00	108.00	44.00	462.00	89.00	51.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	47.00	80.00	21.00
56.00	91.00	103.00	48.00	85.00	138.00	31.00	480.00	101.00	49.00	1.00	13.00	44.00	4.00	49.00	2.00	38.00	1.00	27.00	54.00	3.00
58.00	75.00	88.00	57.00	74.00	101.00	46.00	423.00	91.00	85.00	1.00	35.00	67.00	8.00	57.00	10.00	52.00	1.00	50.00	51.00	2.00
87.00	70.00	89.00	57.00	68.00	93.00	55.00	387.00	77.00	85.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	42.00	64.00	28.00
41.00	100.00	120.00	37.00	90.00	126.00	37.00	535.00	121.00	35.00	1.00	20.00	50.00	1.00	51.00	7.00	50.00	1.00	12.00	51.00	2.00
#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	45.00	64.00	12.00	63.00	22.00	65.00	1.00	13.00	48.00	1.00
50.00	71.00	85.00	60.00	65.00	80.00	57.00	360.00	82.00	82.00	1.00	13.00	44.00	3.00	46.00	6.00	45.00	1.00	13.00	48.00	11.00
50.00	80.00	94.00	54.00	66.00	89.00	57.00	415.00	89.00	57.00	1.00	27.00	54.00	2.00	43.00	10.00	54.00	1.00	36.00	57.00	8.00
#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	32.00	58.00	10.00	56.00	10.00	57.00	1.00	6.00	47.00	4.00
46.00	100.00	125.00	33.00	90.00	130.00	37.00	536.00	122.00	35.00	1.00	18.00	48.00	48.00	55.00	7.00	80.00	1.00	60.00	64.00	6.00
63.00	80.00	80.00	63.00	72.00	96.00	51.00	377.00	79.00	64.00	1.00	8.00	41.00	4.00	47.00	2.00	42.00	1.00	32.00	58.00	16.00
56.00	91.00	103.00	48.00	76.00	103.00	48.00	488.00	103.00	48.00	1.00	16.00	48.00	1.00	39.00	10.00	35.00	1.00	7.00	44.00	7.00
39.00	95.00	108.00	45.00	87.00	132.00	29.00	505.00	114.00	41.00	1.00	17.00	48.00	5.00	50.00	9.00	53.00	1.00	10.00	48.00	4.00
53.00	75.00	89.00	57.00	82.00	85.00	80.00	423.00	91.00	58.00	1.00	7.00	40.00	1.00	37.00	2.00	42.00	1.00	6.00	47.00	2.00
67.00	91.00	103.00	48.00	85.00	116.00	40.00	484.00	107.00	45.00	5.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	74.00	68.00	14.00
53.00	86.00	99.00	61.00	91.00	122.00	35.00	489.00	108.00	45.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	32.00	56.00	.00
56.00	73.00	87.00	59.00	67.00	91.00	56.00	405.00	85.00	59.00	1.00	30.00	55.00	9.00	59.00	12.00	54.00	2.00	#NULL!	#NULL!	#NULL!
71.00	85.00	81.00	63.00	68.00	89.00	57.00	391.00	75.00	67.00	1.00	32.00	57.00	9.00	57.00	14.00	59.00	1.00	8.00	47.00	6.00
51.00	83.00	105.00	47.00	82.00	123.00	35.00	487.00	105.00	45.00	1.00	9.00	40.00	2.00	43.00	2.00	38.00	1.00	15.00	49.00	1.00
57.00	80.00	94.00	64.00	70.00	85.00	53.00	452.00	89.00	51.00	1.00	24.00	52.00	2.00	43.00	13.00	58.00	1.00	27.00	57.00	.00
#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	5.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	6.00	#NULL!	#NULL!	#NULL!
63.00	89.00	84.00	61.00	57.00	81.00	63.00	383.00	80.00	63.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	44.00	60.00	13.00
48.00	100.00	126.00	33.00	78.00	98.00	51.00	472.00	104.00	47.00	1.00	6.00	38.00	4.00	47.00	1.00	38.00	1.00	.00	32.00	.00
43.00	95.00	109.00	45.00	64.00	87.00	69.00	480.00	105.00	45.00	1.00	2.00	29.00	.00	33.00	1.00	37.00	1.00	6.00	45.00	.00
61.00	87.00	99.00	61.00	76.00	103.00	48.00	481.00	101.00	48.00	1.00	28.00	53.00	4.00	49.00	14.00	56.00	1.00	19.00	51.00	.00
37.00	86.00	120.00	37.00	71.00	96.00	53.00	517.00	116.00	39.00	1.00	30.00	57.00	10.00	55.00	12.00	60.00	1.00	.00	32.00	.00
47.00	73.00	87.00	59.00	60.00	82.00	62.00	441.00	85.00	53.00	1.00	28.00	55.00	8.00	55.00	13.00	58.00	1.00	3.00	43.00	1.00
50.00	96.00	111.00	43.00	85.00	114.00	41.00	483.00	107.00	45.00	1.00	12.00	44.00	2.00	43.00	6.00	46.00	1.00	.00	32.00	.00
58.00	87.00	82.00	62.00	71.00	95.00	63.00	404.00	85.00	59.00	1.00	31.00	57.00	8.00	55.00	19.00	63.00	1.00	21.00	52.00	.00
41.00	98.00	106.00	47.00	64.00	87.00	59.00	483.00	109.00	44.00	1.00	20.00	52.00	6.00	50.00	10.00	57.00	1.00	20.00	55.00	1.00
22.00	97.00	113.00	41.00	65.00	114.00	41.00	555.00	127.00	32.00	5.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	6.00	#NULL!	#NULL!	#NULL!
35.00	98.00	117.00	39.00	74.00	101.00	48.00	522.00	117.00	39.00	1.00	24.00	54.00	9.00	55.00	7.00	63.00	1.00	2.00	41.00	.00
43.00	96.00	108.00	45.00	59.00	81.00	63.00	487.00	103.00	48.00	1.00	32.00	57.00	8.00	55.00	11.00	56.00	1.00	4.00	40.00	.00
48.00	79.00	90.00	55.00	67.00	91.00	56.00	423.00	91.00	59.00	1.00	4.00	35.00	3.00	43.00	.00	30.00	1.00	14.00	48.00	9.00
#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	47.00	60.00	2.00
35.00	100.00	125.00	33.00	87.00	117.00	39.00	538.00	122.00	35.00	1.00	74.00	72.00	8.00	55.00	48.00	62.00	1.00	1.00	37.00	.00
47.00	97.00	113.00	41.00	70.00	95.00	53.00	439.00	85.00	53.00	1.00	28.00	58.00	8.00	53.00	16.00	63.00	1.00	52.00	62.00	10.00
59.00	82.00	79.00	64.00	66.00	87.00	65.00	578.00	79.00	64.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	6.00	45.00	4.00
47.00	88.00	104.00	49.00	70.00	95.00	53.00	459.00	101.00	48.00	1.00	23.00	52.00	2.00	43.00	10.00	54.00	1.00	9.00	48.00	7.00
48.00	90.00	94.00	54.00	72.00	96.00	51.00	444.00	87.00	62.00	1.00	28.00	55.00	10.00	58.00	10.00	54.00	1.00	17.00	54.00	14.00
53.00	73.00	87.00	59.00	75.00	103.00	48.00	543.00	95.00	53.00	1.00	15.00	45.00	4.00	49.00	7.00	47.00	1.00	21.00	52.00	6.00
40.00	82.00	104.00	47.00	82.00	111.00	43.00	516.00	116.00	39.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	10.00	46.00	1.00
48.00	82.00	98.00	63.00	76.00	104.00	47.00	488.00	104.00	45.00	1.00	38.00	67.00	14.00	69.00	9.00	90.00	1.00	48.00	60.00	12.00
59.00	83.00	97.00	62.00	79.00	96.00	53.00	431.00	93.00	55.00	1.00	25.00	52.00	.00	34.00	63.00	55.00	1.00	30.00	55.00	1.00
53.00	88.00	84.00	61.00	73.00	108.00	80.00	440.00	95.00	63.00	1.00	3.00	34.00	.00	34.00	1.00	35.00	1.00	1.00	35.00	1.00
39.00	88.00	120.00	37.00	71.00	96.00	53.00	516.00	116.00	39.00	1.00	16.00	47.00	15.00	64.00	13.00	58.00	1.00	11.00	51.00	10.00
48.00	82.00	98.00	63.00	68.00	90.00	57.00	430.00	80.00	55.00	1.00	49.00	65.00	19.00	69.00	13.00	58.00	1.00	30.00	58.00	15.00
55.00	90.00	102.00	49.00	68.00	93.00	55.00	429.00	80.00	55.00	1.00	8.00	48.00	3.00	48.00	2.00	40.00	1.00	13.00	52.00	9.00
58.00	63.00	97.00	62.00	65.00	88.00	57.00	411.00	89.00	58.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	9.00	49.00	1.00
63.00	84.00	108.00	46.00	77.00	105.00	47.00	4													

T3.UNIT3	T3.TOE3	T3.EXT3	PRES.V3	Y3.ITO3	Y3.TT3	Y3.TOM3	Y3.UNIT3	Y3.TOE3	Y3.EXT3	PRES.M3	M3.SO.TO	M3.SO.M3	M3.SO.T3	M3.CO.TO	M3.CO.M3	M3.CO.T3	MS.AF.TO	MS.AF.M3	MS.AF.T3	M3.AC.TO
55.00	8.00	57.00	1.00	31.00	47.00	3.00	37.00	10.00	51.00	1.00	54.00	74.00	87.00	72.00	87.00	82.00	70.00	84.00	84.00	84.00
59.00	2.00	62.00	1.00	29.00	48.00	80.00	48.00	4.00	40.00	1.00	78.00	103.00	48.00	74.00	89.00	51.00	80.00	107.00	48.00	78.00
48.00	8.00	53.00	1.00	35.00	80.00	80.00	51.00	11.00	50.00	1.00	70.00	81.00	55.00	90.00	123.00	35.00	85.00	113.00	41.00	82.00
83.00	1.00	48.00	1.00	#NULL!	40.00	3.00	38.00	18.00	48.00	1.00	68.00	88.00	58.00	82.00	80.00	83.00	88.00	80.00	83.00	85.00
67.00	1.00	48.00	1.00	35.00	80.00	82.00	51.00	12.00	54.00	1.00	73.00	95.00	53.00	88.00	88.00	88.00	75.00	101.00	48.00	85.00
53.00	.00	39.00	1.00	45.00	48.00	9.00	90.00	8.00	48.00	1.00	72.00	94.00	54.00	71.00	84.00	54.00	79.00	106.00	48.00	75.00
51.00	1.00	48.00	1.00	41.00	53.00	8.00	48.00	20.00	61.00	1.00	78.00	105.00	47.00	81.00	109.00	44.00	77.00	104.00	47.00	87.00
88.00	16.00	85.00	1.00	28.00	45.00	12.00	51.00	7.00	48.00	1.00	80.00	80.00	83.00	77.00	103.00	48.00	72.00	87.00	82.00	73.00
57.00	.00	42.00	1.00	73.00	85.00	24.00	82.00	16.00	58.00	1.00	58.00	78.00	85.00	73.00	87.00	53.00	70.00	84.00	84.00	80.00
48.00	9.00	85.00	1.00	71.00	85.00	14.00	58.00	37.00	78.00	1.00	80.00	105.00	46.00	73.00	100.00	50.00	58.00	81.00	83.00	74.00
51.00	2.00	82.00	1.00	17.00	40.00	1.00	32.00	8.00	47.00	1.00	87.00	138.00	25.00	83.00	127.00	32.00	88.00	124.00	34.00	88.00
71.00	.00	39.00	1.00	17.00	40.00	5.00	43.00	5.00	41.00	1.00	71.00	84.00	54.00	89.00	122.00	35.00	82.00	125.00	33.00	87.00
48.00	9.00	84.00	1.00	78.00	40.00	7.00	48.00	7.00	45.00	1.00	90.00	123.00	31.00	79.00	105.00	48.00	86.00	114.00	41.00	78.00
48.00	23.00	83.00	1.00	30.00	41.00	1.00	32.00	8.00	48.00	1.00	88.00	118.00	38.00	88.00	117.00	38.00	80.00	124.00	34.00	77.00
71.00	.00	42.00	1.00	36.00	90.00	12.00	51.00	12.00	54.00	1.00	73.00	95.00	53.00	85.00	84.00	59.00	78.00	101.00	48.00	74.00
48.00	4.00	55.00	1.00	1.00	28.00	.00	26.00	1.00	32.00	1.00	90.00	122.00	35.00	78.00	102.00	48.00	78.00	106.00	47.00	78.00
43.00	3.00	90.00	1.00	54.00	81.00	20.00	63.00	#NULL!	70.00	1.00	75.00	99.00	51.00	89.00	90.00	57.00	72.00	97.00	52.00	73.00
81.00	.00	38.00	1.00	82.00	82.00	18.00	61.00	11.00	50.00	1.00	59.00	78.00	84.00	88.00	88.00	58.00	80.00	83.00	81.00	71.00
59.00	3.00	54.00	1.00	56.00	80.00	19.00	58.00	7.00	48.00	1.00	85.00	85.00	59.00	85.00	85.00	59.00	80.00	83.00	81.00	85.00
51.00	.00	42.00	1.00	37.00	80.00	11.00	49.00	13.00	55.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
54.00	33.00	70.00	1.00	25.00	44.00	4.00	41.00	12.00	52.00	1.00	91.00	124.00	34.00	81.00	108.00	44.00	85.00	128.00	31.00	91.00
85.00	1.00	48.00	1.00	54.00	58.00	19.00	58.00	16.00	58.00	1.00	80.00	108.00	48.00	81.00	108.00	44.00	86.00	90.00	57.00	88.00
56.00	.00	38.00	1.00	39.00	82.00	6.00	45.00	18.00	57.00	1.00	72.00	94.00	54.00	74.00	98.00	51.00	76.00	103.00	48.00	72.00
51.00	8.00	53.00	1.00	#NULL!	55.00	19.00	63.00	17.00	58.00	1.00	79.00	105.00	47.00	81.00	78.00	78.00	71.00	96.00	83.00	82.00
48.00	4.00	55.00	1.00	21.00	42.00	11.00	37.00	8.00	47.00	1.00	79.00	105.00	47.00	78.00	100.00	50.00	74.00	100.00	50.00	71.00
84.00	27.00	85.00	1.00	88.00	85.00	22.00	85.00	24.00	85.00	1.00	87.00	117.00	36.00	70.00	92.00	55.00	88.00	116.00	38.00	81.00
38.00	18.00	81.00	1.00	#NULL!	48.00	8.00	45.00	10.00	48.00	1.00	90.00	122.00	35.00	70.00	92.00	55.00	82.00	122.00	35.00	85.00
#NULL!	#NULL!	#NULL!	1.00	83.00	83.00	16.00	61.00	18.00	58.00	1.00	83.00	83.00	81.00	82.00	80.00	83.00	82.00	85.00	80.00	85.00
53.00	.00	42.00	1.00	#NULL!	88.00	32.00	89.00	18.00	59.00	1.00	85.00	85.00	80.00	87.00	87.00	59.00	80.00	83.00	87.00	83.00
43.00	8.00	53.00	1.00	45.00	53.00	4.00	41.00	18.00	59.00	1.00	84.00	84.00	81.00	85.00	86.00	59.00	86.00	90.00	57.00	89.00
37.00	29.00	67.00	1.00	30.00	47.00	2.00	35.00	13.00	55.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
#NULL!	#NULL!	#NULL!	5.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	5.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
83.00	16.00	81.00	1.00	76.00	88.00	23.00	88.00	30.00	72.00	1.00	76.00	100.00	80.00	89.00	90.00	57.00	88.00	92.00	56.00	88.00
37.00	.00	42.00	1.00	17.00	40.00	11.00	49.00	1.00	32.00	1.00	87.00	117.00	39.00	88.00	121.00	36.00	91.00	121.00	38.00	86.00
37.00	5.00	58.00	1.00	51.00	58.00	13.00	52.00	16.00	59.00	1.00	75.00	99.00	51.00	70.00	92.00	55.00	88.00	93.00	55.00	82.00
36.00	10.00	57.00	1.00	41.00	53.00	11.00	53.00	14.00	54.00	1.00	88.00	118.00	38.00	88.00	121.00	36.00	82.00	108.00	44.00	72.00
37.00	.00	42.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	1.00	83.00	111.00	43.00	84.00	114.00	41.00	93.00	124.00	34.00	84.00
43.00	.00	42.00	1.00	15.00	39.00	8.00	45.00	1.00	32.00	1.00	83.00	111.00	43.00	74.00	98.00	51.00	71.00	96.00	53.00	84.00
37.00	.00	42.00	1.00	30.00	42.00	9.00	47.00	2.00	35.00	1.00	82.00	109.00	44.00	74.00	98.00	51.00	74.00	100.00	50.00	83.00
38.00	11.00	88.00	1.00	78.00	85.00	18.00	61.00	33.00	84.00	1.00	84.00	84.00	81.00	83.00	81.00	83.00	88.00	90.00	57.00	88.00
43.00	12.00	82.00	7.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	7.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
#NULL!	#NULL!	#NULL!	5.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	5.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
37.00	2.00	82.00	1.00	37.00	80.00	15.00	55.00	11.00	52.00	1.00	82.00	108.00	44.00	78.00	105.00	47.00	80.00	107.00	45.00	89.00
36.00	.00	38.00	1.00	45.00	55.00	11.00	53.00	16.00	57.00	1.00	80.00	108.00	48.00	88.00	121.00	38.00	98.00	121.00	38.00	84.00
58.00	.00	38.00	1.00	53.00	88.00	20.00	83.00	14.00	54.00	1.00	89.00	90.00	57.00	71.00	94.00	54.00	70.00	94.00	54.00	73.00
48.00	29.00	85.00	1.00	16.00	80.00	13.00	55.00	19.00	60.00	1.00	72.00	94.00	54.00	87.00	87.00	58.00	81.00	84.00	81.00	80.00
37.00	.00	42.00	1.00	38.00	80.00	13.00	52.00	9.00	49.00	1.00	89.00	120.00	37.00	72.00	96.00	53.00	76.00	103.00	48.00	84.00
80.00	12.00	88.00	1.00	22.00	43.00	8.00	48.00	5.00	41.00	1.00	88.00	118.00	38.00	76.00	102.00	49.00	87.00	115.00	40.00	85.00
51.00	.00	42.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
56.00	1.00	48.00	1.00	38.00	50.00	13.00	52.00	14.00	55.00	1.00	70.00	91.00	56.00	75.00	100.00	50.00	81.00	108.00	45.00	77.00
84.00	2.00	82.00	1.00	16.00	40.00	7.00	44.00	8.00	47.00	1.00	77.00	102.00	48.00	73.00	97.00	52.00	75.00	101.00	48.00	75.00
54.00	9.00	58.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
43.00	4.00	51.00	1.00	57.00	84.00	9.00	50.00	25.00	87.00	1.00	78.00	100.00	50.00	77.00	103.00	48.00	83.00	110.00	43.00	81.00
82.00	24.00	84.00	1.00	44.00	54.00	9.00	50.00	16.00	57.00	1.00	79.00	105.00	47.00	74.00	99.00	51.00	71.00	96.00	53.00	89.00
43.00	17.00	80.00	1.00	48.00	55.00	4.00	41.00	22.00	83.00	1.00	89.00	90.00	87.00	86.00	88.00	58.00	88.00	92.00	55.00	85.00
43.00	.00	39.00	1.00	24.00	44.00	5.00	45.00	9.00	48.00	1.00	79.00	105.00	47.00	77.00	103.00	48.00	78.00	103.00	48.00	75.00
80.00	.00	42.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
84.00	2.00	82.00	1.00	16.00	39.00	9.00	47.00	2.00	35.00	1.00	70.00	91.00	56.00	89.00	90.00	57.00	72.00	97.00	52.00	75.00
89.00	1.00	49.00	1.00	10.00	36.00	2.00	35.00	2.00	35.00	1.00	75.00	98.00	51.00	72.00	95.00	53.00	71.00	96.00	53.00	71.00
43.00	8.00	89.00	1.00	67.00	83.00	27.00	85.00	18.00	82.00	2.00	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!
43.00	30.00	88.00	1.00	148.00	87.00	45.00	85.0													

[illegible]

TS.TOW	TS.INTS	TS.TOX	TS.EXTS	PRIS.YS	YS.TTO	YS.TTS	YS.TOW	YS.INTS	YS.TOX	YS.EXTS	PRIS.MS	MS.SQ.TO	MS.SQ.MS	MS.SQ.TS	MS.CO.TO	MS.CO.MS	MS.CO.TS	MS.AF.TO	MS.AF.MS	MS.AF.TS
2.00	40.00	.00	42.00	1.00	20.00	42.00	3.00	37.00	8.00	47.00	1.00	85.00	114.00	41.00	81.00	108.00	44.00	80.00	118.00	37.00
5.00	55.00	.00	42.00	1.00	32.00	48.00	8.00	47.00	18.00	51.00	1.00	77.00	102.00	49.00	75.00	100.00	50.00	78.00	108.00	48.00
3.00	45.00	1.00	42.00	1.00	25.00	44.00	4.00	41.00	11.00	50.00	1.00	79.00	106.00	47.00	82.00	111.00	43.00	83.00	110.00	43.00
10.00	60.00	1.00	48.00	1.00	51.00	60.00	13.00	55.00	9.00	46.00	1.00	88.00	90.00	57.00	89.00	90.00	67.00	73.00	98.00	51.00
2.00	48.00	2.00	48.00	1.00	61.00	61.00	13.00	62.00	20.00	64.00	1.00	88.00	84.00	58.00	78.00	102.00	48.00	81.00	121.00	38.00
8.00	65.00	.00	40.00	1.00	31.00	48.00	11.00	53.00	7.00	45.00	1.00	70.00	91.00	56.00	79.00	108.00	48.00	75.00	101.00	48.00
1.00	43.00	.00	30.00	1.00	32.00	50.00	8.00	50.00	20.00	61.00	1.00	86.00	114.00	41.00	81.00	108.00	44.00	88.00	116.00	39.00
4.00	53.00	1.00	46.00	1.00	38.00	60.00	15.00	54.00	6.00	44.00	1.00	53.00	73.00	68.00	85.00	84.00	61.00	84.00	87.00	58.00
4.00	51.00	.00	43.00	1.00	40.00	62.00	18.00	58.00	4.00	40.00	1.00	59.00	79.00	64.00	86.00	115.00	40.00	77.00	103.00	48.00
9.00	59.00	25.00	65.00	1.00	104.00	75.00	23.00	65.00	56.00	94.00	1.00	92.00	125.00	33.00	84.00	114.00	41.00	81.00	84.00	81.00
.00	37.00	.00	42.00	1.00	17.00	49.00	1.00	32.00	7.00	45.00	1.00	98.00	135.00	27.00	103.00	146.00	20.00	85.00	137.00	25.00
NULL	NULL	NULL	NULL	7.00	NULL	NULL	NULL	NULL	NULL	NULL	2.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
.00	35.00	7.00	54.00	1.00	30.00	42.00	3.00	37.00	8.00	47.00	1.00	85.00	114.00	41.00	81.00	108.00	44.00	80.00	118.00	37.00
3.00	51.00	37.00	60.00	1.00	89.00	72.00	30.00	72.00	22.00	63.00	1.00	98.00	141.00	23.00	85.00	115.00	40.00	81.00	121.00	36.00
4.00	53.00	14.00	64.00	1.00	61.00	61.00	63.00	52.00	20.00	64.00	1.00	72.00	94.00	54.00	68.00	86.00	69.00	77.00	104.00	47.00
3.00	51.00	1.00	48.00	1.00	129.00	65.00	41.00	78.00	38.00	82.00	1.00	92.00	128.00	33.00	85.00	84.00	81.00	88.00	90.00	57.00
15.00	65.00	18.00	62.00	1.00	30.00	47.00	19.00	51.00	18.00	48.00	1.00	73.00	96.00	53.00	74.00	99.00	51.00	78.00	103.00	48.00
1.00	45.00	.00	40.00	1.00	41.00	53.00	12.00	54.00	9.00	48.00	1.00	56.00	75.00	68.00	87.00	87.00	79.00	85.00	80.00	57.00
4.00	53.00	1.00	48.00	1.00	54.00	66.00	11.00	48.00	17.00	61.00	1.00	77.00	102.00	49.00	67.00	87.00	89.00	82.00	83.00	81.00
5.00	56.00	3.00	54.00	1.00	79.00	67.00	22.00	60.00	27.00	71.00	1.00	71.00	85.00	55.00	55.00	89.00	71.00	89.00	82.00	55.00
6.00	57.00	31.00	67.00	1.00	19.00	49.00	4.00	41.00	8.00	48.00	1.00	91.00	124.00	34.00	91.00	124.00	34.00	97.00	134.00	27.00
15.00	64.00	.00	42.00	1.00	90.00	66.00	19.00	58.00	15.00	48.00	1.00	83.00	111.00	43.00	75.00	100.00	50.00	86.00	114.00	41.00
7.00	58.00	.00	40.00	1.00	55.00	69.00	8.00	45.00	23.00	64.00	1.00	82.00	109.00	44.00	82.00	90.00	83.00	81.00	108.00	46.00
4.00	53.00	14.00	60.00	1.00	37.00	61.00	16.00	59.00	9.00	48.00	1.00	89.00	120.00	37.00	80.00	108.00	45.00	77.00	133.00	NULL
3.00	51.00	8.00	59.00	1.00	9.00	35.00	3.00	37.00	5.00	42.00	1.00	81.00	108.00	45.00	78.00	105.00	47.00	75.00	101.00	48.00
NULL	NULL	NULL	NULL	5.00	NULL	NULL	NULL	NULL	NULL	NULL	5.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3.00	51.00	28.00	67.00	1.00	69.00	62.00	7.00	45.00	26.00	70.00	1.00	97.00	138.00	25.00	80.00	108.00	45.00	80.00	120.00	37.00
2.00	48.00	.00	40.00	1.00	37.00	61.00	13.00	55.00	8.00	48.00	1.00	79.00	94.00	54.00	63.00	81.00	63.00	65.00	86.00	58.00
5.00	55.00	.00	42.00	1.00	65.00	62.00	21.00	59.00	21.00	65.00	1.00	81.00	108.00	45.00	68.00	88.00	59.00	88.00	92.00	56.00
2.00	48.00	4.00	52.00	1.00	9.00	33.00	1.00	32.00	2.00	33.00	1.00	86.00	115.00	40.00	80.00	108.00	45.00	77.00	104.00	47.00
6.00	56.00	33.00	72.00	1.00	NULL	45.00	2.00	35.00	8.00	42.00	1.00	81.00	124.00	34.00	87.00	87.00	59.00	71.00	98.00	53.00
NULL	NULL	NULL	NULL	5.00	NULL	NULL	NULL	NULL	NULL	NULL	5.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
18.00	68.00	53.00	62.00	2.00	NULL	NULL	NULL	NULL	NULL	NULL	.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
3.00	51.00	.00	42.00	1.00	9.00	34.00	8.00	41.00	1.00	32.00	1.00	81.00	108.00	45.00	74.00	99.00	51.00	77.00	104.00	47.00
1.00	45.00	14.00	64.00	1.00	65.00	62.00	18.00	55.00	25.00	69.00	1.00	85.00	114.00	41.00	70.00	92.00	55.00	72.00	97.00	52.00
2.00	48.00	27.00	66.00	1.00	25.00	44.00	8.00	48.00	10.00	49.00	1.00	96.00	135.00	27.00	85.00	115.00	40.00	91.00	121.00	36.00
3.00	51.00	14.00	64.00	1.00	19.00	41.00	4.00	39.00	7.00	45.00	1.00	92.00	126.00	33.00	89.00	120.00	37.00	93.00	124.00	34.00
1.00	45.00	.00	42.00	1.00	39.00	61.00	20.00	58.00	8.00	44.00	1.00	79.00	105.00	47.00	81.00	108.00	44.00	74.00	100.00	50.00
.00	38.00	.00	42.00	1.00	19.00	41.00	19.00	48.00	3.00	38.00	1.00	81.00	108.00	45.00	78.00	105.00	47.00	80.00	107.00	48.00
1.00	45.00	33.00	68.00	2.00	NULL	NULL	NULL	NULL	NULL	NULL	2.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
NULL	NULL	NULL	NULL	7.00	NULL	NULL	NULL	NULL	NULL	NULL	2.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
NULL	NULL	NULL	NULL	5.00	NULL	NULL	NULL	NULL	NULL	NULL	2.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
1.00	45.00	3.00	54.00	1.00	41.00	52.00	17.00	55.00	15.00	55.00	1.00	79.00	105.00	47.00	77.00	103.00	48.00	74.00	100.00	50.00
2.00	48.00	.00	40.00	1.00	35.00	44.00	5.00	43.00	9.00	48.00	1.00	88.00	120.00	37.00	80.00	123.00	35.00	93.00	124.00	34.00
3.00	51.00	.00	48.00	1.00	30.00	41.00	3.00	38.00	9.00	48.00	1.00	78.00	100.00	50.00	70.00	92.00	55.00	74.00	100.00	50.00
NULL	NULL	NULL	NULL	5.00	NULL	NULL	NULL	NULL	NULL	NULL	2.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
1.00	45.00	7.00	60.00	1.00	21.00	42.00	9.00	47.00	2.00	35.00	1.00	83.00	128.00	31.00	85.00	118.00	40.00	81.00	121.00	36.00
7.00	60.00	19.00	62.00	1.00	62.00	73.00	30.00	72.00	28.00	71.00	1.00	43.00	64.00	74.00	57.00	72.00	69.00	51.00	75.00	67.00
13.00	62.00	.00	42.00	1.00	23.00	43.00	8.00	43.00	9.00	49.00	1.00	84.00	84.00	61.00	59.00	75.00	67.00	68.00	86.00	59.00
.00	35.00	.00	42.00	1.00	61.00	61.00	16.00	57.00	20.00	64.00	1.00	87.00	87.00	59.00	78.00	102.00	49.00	80.00	107.00	48.00
.00	37.00	.00	42.00	1.00	7.00	32.00	3.00	37.00	2.00	35.00	1.00	78.00	103.00	48.00	73.00	97.00	52.00	74.00	100.00	50.00
.00	38.00	2.00	48.00	3.00	NULL	NULL	NULL	NULL	NULL	NULL	3.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
.00	38.00	7.00	55.00	1.00	43.00	54.00	8.00	45.00	21.00	62.00	1.00	91.00	124.00	34.00	80.00	108.00	45.00	89.00	118.00	38.00
.00	38.00	4.00	52.00	1.00	15.00	35.00	4.00	41.00	1.00	30.00	1.00	93.00	128.00	31.00	85.00	115.00	40.00	92.00	122.00	36.00
.00	38.00	14.00	59.00	1.00	48.00	56.00	9.00	50.00	17.00	58.00	1.00	72.00	94.00	54.00	67.00	87.00	68.00	93.00	124.00	34.00
.00	38.00	.00	38.00	1.00	32.00	46.00	9.00	48.00	16.00	57.00	1.00	78.00	103.00	48.00	76.00	102.00	48.00	74.00	100.00	50.00
.00	37.00	.00	42.00	3.00	NULL	NULL	NULL	NULL	NULL	NULL	3.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
.00	37.00	.00	42.00	1.00	38.00	7.00	44.00	.00	27.00	.00	1.00	74.00	97.00	52.00	86.00	117.00	39.00	88.00	116.00	38.00
1.00	45.00	1.00	48.00	1.00	23.00	43.00	12.00	61.00	3.00	39.00	1.00	70.00	91.00	56.00	72.00	96.00	53.00	71.00	96.00	53.00
NULL	NULL	NULL	NULL	7.00	NULL	NULL	NULL	NULL	NULL	NULL	7.00	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL	NULL
2.00	48.00	34.00	68.00	1.00	185.00	93.00	80.00	90.00	80.00	89.00	1.00	71.00	83.00	58.00	83.00	83.00	58.00	83.00	79.00	84.00
.00	36.00	.00	40.00	1.00	67.00	65.00	21.00	64.00	12.00	62.00	1.00	63.00	83.00							

	MS.AC.MS	MS.AC.TS	MS.FA.TO	MS.FA.ME	MS.FA.TS	MS.PH.TO	MS.PH.MS	MS.PH.TS	MS.TY.TO	MS.TY.MS	MS.TY.TS	PRESC23S	PREST23S	PREST23S	PREMA23S	ALL	A1C	A2I	A2C	A3.SI
85.00	116.00	48.00	88.80	106.00	47.80	85.00	114.00	41.00	518.00	116.00	39.00	1.00	1.00	1.00	1.00	83.80	84.10	100.00	84.03	100.00
75.00	101.00	48.00	76.00	88.00	87.00	75.00	103.00	48.00	453.00	108.00	80.00	.00	1.00	1.00	1.00	75.83	84.10	85.24	84.03	85.39
87.00	134.00	27.00	88.80	117.00	29.00	79.00	85.00	53.00	508.00	114.00	41.00	1.00	1.00	1.00	1.00	86.18	84.10	80.86	84.03	80.83
74.00	102.00	48.00	83.00	87.80	82.80	87.00	91.00	88.00	438.00	84.00	54.00	.00	1.00	1.00	1.00	80.74	84.10	87.82	84.03	80.41
83.00	113.00	41.00	100.80	125.00	33.00	71.00	86.00	83.00	488.00	108.00	46.00	1.00	1.00	.00	.00	86.30	84.10	85.24	84.03	80.41
70.00	85.00	53.00	84.00	108.80	46.00	89.00	84.00	54.00	467.00	100.00	80.00	1.00	1.00	1.00	1.00	85.19	84.10	85.24	84.03	86.83
83.00	116.00	43.00	84.00	106.80	46.00	84.00	113.00	41.00	516.00	115.00	46.00	1.00	1.00	1.00	1.00	100.00	84.10	86.41	84.03	100.00
83.00	84.00	81.00	78.00	82.80	85.00	89.00	80.00	57.00	388.00	82.00	82.00	1.00	1.00	1.00	1.00	86.30	84.10	84.44	84.03	82.88
86.00	110.00	37.00	108.00	135.80	33.00	85.00	114.00	41.00	484.00	110.00	43.00	1.00	1.00	.00	.00	88.89	84.10	83.85	84.03	85.89
33.00	84.00	81.00	83.00	87.80	82.00	83.00	112.00	42.00	436.00	86.00	53.00	1.00	1.00	1.00	1.00	86.37	84.10	87.82	84.03	86.83
85.00	129.00	31.00	108.00	135.80	33.00	84.00	128.00	33.00	583.00	140.00	23.00	.00	1.00	1.00	1.00	88.16	84.10	100.00	84.03	100.00
84.00	116.00	48.00	83.00	108.80	47.80	85.00	114.00	41.00	518.00	116.00	38.00	.00	1.00	1.00	1.00	100.00	84.10	86.41	84.03	86.83
74.00	102.00	48.00	78.00	82.80	85.80	79.00	107.00	45.00	304.00	113.00	41.00	.00	1.00	1.00	1.00	80.74	84.10	86.83	84.03	87.28
83.00	84.00	81.00	84.00	108.80	48.00	86.00	80.00	57.00	438.00	85.00	53.00	.00	1.00	1.00	1.00	85.19	84.10	84.44	84.03	85.89
59.00	80.00	83.00	72.00	86.00	88.00	87.00	91.00	88.00	421.00	81.00	56.00	.00	1.00	1.00	1.00	88.89	84.10	88.87	84.03	88.18
73.00	104.00	47.00	78.00	82.00	87.80	78.00	104.00	47.00	450.00	88.00	51.00	.00	1.00	1.00	.00	88.30	84.10	100.00	84.03	88.83
71.00	95.00	53.00	78.00	85.00	88.00	58.00	78.00	85.00	386.00	81.00	63.00	1.00	1.00	1.00	1.00	84.00	84.10	84.00	84.03	100.00
74.00	102.00	48.00	88.00	118.00	43.00	75.00	103.00	46.00	440.00	88.00	61.00	1.00	1.00	1.00	1.00	100.00	81.94	88.18	83.05	86.32
80.00	81.00	88.00	80.00	84.00	54.00	80.00	82.0													

	A23C	A3C	A4	A5C	A5.1	A5.1C	A5.2	A5.2C	B1ATV	B1ATX	B1ATPER	B1BTN	B1BTCH	B2ATV	B2ATX	B2ATPER	B2BTCH	B3ATV	B3ATX	B3ATPER
05.71	100.00	06.52	06.28	06.70	01.53	05.10	100.00	04.24	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	28.00	3.00	.84	#NULL!	28.00	3.00	.80
06.71	92.50	06.52	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	28.00	3.00	.80	#NULL!	28.00	2.00	.93
06.71	100.00	06.52	100.00	04.52	100.00	03.10	100.00	05.61	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	28.00	2.00	.80	#NULL!	24.00	6.00	.90
06.71	06.10	06.52	06.28	04.52	01.26	05.10	06.70	06.61	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	28.00	1.00	.98	#NULL!	28.00	4.00	.84
06.71	100.00	06.52	06.58	04.52	00.47	03.10	02.06	06.61	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	25.00	6.00	.81	#NULL!	24.00	6.00	.80
06.71	100.00	06.52	04.38	02.90	03.22	04.23	04.12	06.32	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	25.00	4.00	.83	#NULL!	27.00	3.00	.90
06.71	100.00	06.52	00.52	02.90	00.31	04.33	100.00	06.32	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	27.00	6.00	.84	#NULL!	25.00	4.00	.88
06.71	75.90	06.52	03.70	06.70	03.05	03.10	06.70	04.24	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	22.00	5.00	.98	#NULL!	25.00	6.00	.81
06.71	00.74	06.52	02.70	02.44	00.83	03.91	00.71	03.25	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	23.00	2.00	.92	#NULL!	30.00	2.00	.94
06.71	00.74	06.52	01.30	02.90	00.10	03.43	00.53	06.76	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	30.00	2.00	.94	#NULL!	30.00	1.00	.97
06.71	06.16	06.52	06.28	02.44	04.07	03.91	07.70	03.25	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	28.00	1.00	.97	#NULL!	23.00	2.00	.94
#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	28.00	2.00	.80	#NULL!	28.00	3.00	.90
06.71	06.16	06.52	00.52	02.90	77.97	04.33	00.71	06.32	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	27.00	4.00	.87	#NULL!	26.00	3.00	.87
06.71	00.74	06.52	04.48	06.30	01.53	03.43	00.44	06.76	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	15.00	6.00	.48	#NULL!	22.00	0.00	.71
06.71	04.44	06.52	00.10	02.44	00.22	03.91	77.94	03.25	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	27.00	2.00	.87	#NULL!	28.00	3.00	.90
06.71	01.40	06.52	01.90	06.70	00.22	03.10	#NULL!	04.24	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	31.00	1.00	.97	#NULL!	20.00	0.00	.71
06.71	100.00	06.52	06.40	06.30	100.00	03.43	06.76	06.76	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	04.00	6.00	.81	#NULL!	28.00	2.00	.93
02.17	100.00	03.82	100.00	03.83	02.50	00.20	00.56	00.50	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	24.00	6.00	.80	100.00	27.00	6.00	.84
02.17	100.00	03.82	100.00	03.83	06.67	00.20	00.56	03.56	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	20.00	9.00	.58	100.00	27.00	4.00	.87
02.17	02.20	03.82	00.06	03.83	00.17	00.20	01.30	06.22	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	29.00	1.00	.87	100.00	32.00	.00	1.00
02.17	00.06	03.82	00.26	03.83	00.83	00.20	03.46	01.20	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	28.00	4.00	.87	100.00	19.00	11.00	.51
02.17	06.28	03.82	100.00	03.83	06.67	00.20	07.83	00.04	#NULL!	#NULL!	#NULL!	#NULL!	#NULL!	24.00						

[illegible]

[illegible]

Appendix 2: Data Collection Instruments

- Child Behaviour Checklist
- Youth Self report
- Teacher Report Form
- Multi-Dimensional Self Concept Scale

BEST COPY

AVAILABLE

TEXT IN ORIGINAL IS
CLOSE TO THE EDGE OF
THE PAGE

CHILD BEHAVIOR CHECKLIST FOR AGES 4-18

For office use only
ID #

Please Print

CHILD'S FULL NAME	FIRST	MIDDLE	LAST	PARENTS' USUAL TYPE OF WORK, even if not working now. (Please be specific—for example, auto mechanic, high school teacher, homemaker, laborer, lathe operator, shoe salesman, army sergeant.)
SEX	AGE		ETHNIC GROUP OR RACE	
TODAY'S DATE		CHILD'S BIRTHDATE		
GRADE IN SCHOOL		Please fill out this form to reflect <i>your</i> view of the child's behavior even if other people might not agree. Feel free to print additional comments beside each item and in the spaces provided on page 2.		
NOT ATTENDING SCHOOL				FATHER'S TYPE OF WORK:
				MOTHER'S TYPE OF WORK:
				THIS FORM FILLED OUT BY:
				<input type="checkbox"/> Mother (full name)
				<input type="checkbox"/> Father (full name)
				<input type="checkbox"/> Other—full name & relationship to child:

I. Please list the sports your child most likes to take part in. For example: swimming, baseball, skating, skate boarding, bike riding, fishing, etc.	Compared to others of the same age, about how much time does he/she spend in each?	Compared to others of the same age, how well does he/she do each one?
<input type="checkbox"/> None	Don't Know Less Than Average Average More Than Average	Don't Know Below Average Average Above Average
a. <u>SWIMMING</u>	<input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>
b. <u>BASKETBALL</u>	<input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>
c. _____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
II. Please list your child's favorite hobbies, activities, and games, other than sports. For example: stamps, dolls, books, piano, crafts, cars, singing, etc. (Do <i>not</i> include listening to radio or TV.)	Compared to others of the same age, about how much time does he/she spend in each?	Compared to others of the same age, how well does he/she do each one?
<input type="checkbox"/> None	Don't Know Less Than Average Average More Than Average	Don't Know Below Average Average Above Average
a. <u>READING</u>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/>
b. _____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
c. _____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
III. Please list any organizations, clubs, teams, or groups your child belongs to.	Compared to others of the same age, how active is he/she in each?	
<input checked="" type="checkbox"/> None	Don't Know Less Active Average More Active	
a. _____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
b. _____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
c. _____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
IV. Please list any jobs or chores your child has. For example: paper route, babysitting, making bed, working in store, etc. (Include <i>both</i> paid and unpaid jobs and chores.)	Compared to others of the same age, how well does he/she carry them out?	
<input type="checkbox"/> None	Don't Know Below Average Average Above Average	
a. <u>TIDIES ROOM</u>	<input type="checkbox"/> <input type="checkbox"/> <input checked="" type="checkbox"/> <input type="checkbox"/>	
b. _____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	
c. _____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>	

Please Print

1. About how many close friends does your child have? ☐ None ☐ 1 ☒ 2 or 3 ☐ 4 or more
(Do not include brothers & sisters)

2. About how many times a week does your child do things with any friends outside of regular school hours?
(Do not include brothers & sisters) ☐ Less than 1 ☐ 1 or 2 ☒ 3 or more

VI. Compared to others of his/her age, how well does your child:

	Worse	About Average	Better	
a. Get along with his/her brothers & sisters?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/> Has no brothers or sisters
b. Get along with other kids?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
c. Behave with his/her parents?	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	
d. Play and work alone?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	

VII. 1. For ages 6 and older—performance in academic subjects. ☐ Does not attend school because _____

Check a box for each subject that child takes

	Failing	Below Average	Average	Above Average
a. Reading, English, or Language Arts	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
b. History or Social Studies	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
c. Arithmetic or Math	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
d. Science	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
e. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Other academic subjects—for example: computer courses, foreign language, business. Do not include gym, shop, driver's ed., etc.

2. Does your child receive special remedial services or attend a special class or special school? ☒ No ☐ Yes—kind of services, class, or school: _____

3. Has your child repeated any grades? ☒ No ☐ Yes—grades and reasons: _____

4. Has your child had any academic or other problems in school? ☒ No ☐ Yes—please describe: _____

When did these problems start? _____

Have these problems ended? ☐ No ☐ Yes—when? _____

Does your child have any illness or disability (either physical or mental)? ☒ No ☐ Yes—please describe: _____

What concerns you most about your child? NOTHING

Please describe the best things about your child: WELL BEHAVED, TRUSTING.

Below is a list of items that describe children and youth. For each item that describes your child **now or within the past 6 months**, please circle the **2** if the item is **very true or often true** of your child. Circle the **1** if the item is **somewhat or sometimes true** of your child. If the item is **not true** of your child, circle the **0**. Please answer all items as well as you can, even if some do not seem to apply to your child.

Please Print

0 = Not True (as far as you know) 1 = Somewhat or Sometimes True 2 = Very True or Often True

0	1	2	1.	Acts too young for his/her age	0	1	2	31.	Fears he/she might think or do something bad
0	1	2	2.	Allergy (describe): <u>Penicillin</u>	0	1	2	32.	Feels he/she has to be perfect
0	1	2	3.	Argues a lot	0	1	2	33.	Feels or complains that no one loves him/her
0	1	2	4.	Asthma	0	1	2	34.	Feels others are out to get him/her
0	1	2	5.	Behaves like opposite sex	0	1	2	35.	Feels worthless or inferior
0	1	2	6.	Bowel movements outside toilet	0	1	2	36.	Gets hurt a lot, accident-prone
0	1	2	7.	Bragging, boasting	0	1	2	37.	Gets in many fights
0	1	2	8.	Can't concentrate, can't pay attention for long	0	1	2	38.	Gets teased a lot
0	1	2	9.	Can't get his/her mind off certain thoughts; obsessions (describe): _____	0	1	2	39.	Hangs around with others who get in trouble
0	1	2	10.	Can't sit still, restless, or hyperactive	0	1	2	40.	Hears sounds or voices that aren't there (describe): _____
0	1	2	11.	Clings to adults or too dependent	0	1	2	41.	Impulsive or acts without thinking
0	1	2	12.	Complains of loneliness	0	1	2	42.	Would rather be alone than with others
0	1	2	13.	Confused or seems to be in a fog	0	1	2	43.	Lying or cheating
0	1	2	14.	Cries a lot	0	1	2	44.	Bites fingernails
0	1	2	15.	Cruel to animals	0	1	2	45.	Nervous, highstrung, or tense
0	1	2	16.	Cruelty, bullying, or meanness to others	0	1	2	46.	Nervous movements or twitching (describe): _____
0	1	2	17.	Day-dreams or gets lost in his/her thoughts	0	1	2	47.	Nightmares
0	1	2	18.	Deliberately harms self or attempts suicide	0	1	2	48.	Not liked by other kids
0	1	2	19.	Demands a lot of attention	0	1	2	49.	Constipated, doesn't move bowels.
0	1	2	20.	Destroys his/her own things	0	1	2	50.	Too fearful or anxious
0	1	2	21.	Destroys things belonging to his/her family or others	0	1	2	51.	Feels dizzy
0	1	2	22.	Disobedient at home	0	1	2	52.	Feels too guilty
0	1	2	23.	Disobedient at school	0	1	2	53.	Overeating
0	1	2	24.	Doesn't eat well	0	1	2	54.	Overtired
0	1	2	25.	Doesn't get along with other kids	0	1	2	55.	Overweight
0	1	2	26.	Doesn't seem to feel guilty after misbehaving	0	1	2	56.	Physical problems without known medical cause:
0	1	2	27.	Easily jealous	0	1	2	a.	Aches or pains (not stomach or headaches)
0	1	2	28.	Eats or drinks things that are not food — don't include sweets (describe): _____	0	1	2	b.	Headaches
0	1	2	29.	Fears certain animals, situations, or places, other than school (describe): <u>Fears</u> <u>dogs & wasps.</u>	0	1	2	c.	Nausea, feels sick
0	1	2	30.	Fears going to school	0	1	2	d.	Problems with eyes (not if corrected by glasses) (describe): _____
					0	1	2	e.	Rashes or other skin problems
					0	1	2	f.	Stomachaches or cramps
					0	1	2	g.	Vomiting, throwing up
					0	1	2	h.	Other (describe): _____

Please Print

0 = Not True (as far as you know)

1 = Somewhat or Sometimes True

2 = Very True or Often True

1	2	57.	Physically attacks people	0	1	2	84.	Strange behavior (describe):
1	2	58.	Picks nose, skin, or other parts of body (describe):					
				0	1	2	85.	Strange ideas (describe):
1	2	59.	Plays with own sex parts in public					
1	2	60.	Plays with own sex parts too much	0	1	2	86.	Stubborn, sullen, or irritable
1	2	61.	Poor school work	0	1	2	87.	Sudden changes in mood or feelings
1	2	62.	Poorly coordinated or clumsy	0	1	2	88.	Sulks a lot
1	2	63.	Prefers being with older kids	0	1	2	89.	Suspicious
1	2	64.	Prefers being with younger kids	0	1	2	90.	Swearing or obscene language
1	2	65.	Refuses to talk	0	1	2	91.	Talks about killing self
1	2	66.	Repeats certain acts over and over; compulsions (describe):	0	1	2	92.	Talks or walks in sleep (describe):
1	2	67.	Runs away from home	0	1	2	93.	Talks too much
1	2	68.	Screams a lot	0	1	2	94.	Teases a lot
1	2	69.	Secretive, keeps things to self	0	1	2	95.	Temper tantrums or hot temper
1	2	70.	Sees things that aren't there (describe):	0	1	2	96.	Thinks about sex too much
				0	1	2	97.	Threatens people
				0	1	2	98.	Thumb-sucking
				0	1	2	99.	Too concerned with neatness or cleanliness
1	2	71.	Self-conscious or easily embarrassed	0	1	2	100.	Trouble sleeping (describe):
1	2	72.	Sets fires					
1	2	73.	Sexual problems (describe):	0	1	2	101.	Truancy, skips school
				0	1	2	102.	Underactive, slow moving, or lacks energy
				0	1	2	103.	Unhappy, sad, or depressed
1	2	74.	Showing off or clowning	0	1	2	104.	Unusually loud
1	2	75.	Shy or timid	0	1	2	105.	Uses alcohol or drugs for nonmedical purposes (describe):
1	2	76.	Sleeps less than most kids	0	1	2	106.	Vandalism
1	2	77.	Sleeps more than most kids during day and/or night (describe):	0	1	2	107.	Wets self during the day
				0	1	2	108.	Wets the bed
1	2	78.	Smears or plays with bowel movements	0	1	2	109.	Whining
1	2	79.	Speech problem (describe):	0	1	2	110.	Wishes to be of opposite sex
				0	1	2	111.	Withdrawn, doesn't get involved with others
1	2	80.	Stares blankly	0	1	2	112.	Worries
1	2	81.	Steals at home				113.	Please write in any problems your child has that were not listed above:
1	2	82.	Steals outside the home					
1	2	83.	Stores up things he/she doesn't need (describe):	0	1	2		
				0	1	2		
				0	1	2		

PLEASE BE SURE YOU HAVE ANSWERED ALL ITEMS.

UNDERLINE ANY YOU ARE CONCERNED ABOUT.

YOUTH SELF-REPORT FOR AGES 11-18

For office use only
ID #

Please Print

YOUR FULL NAME FIRST MIDDLE LAST

YOUR SEX

☒ Boy ☐ Girl

YOUR AGE

ETHNIC GROUP OR RACE

PARENTS' USUAL TYPE OF WORK, even if not working now (Please be specific—for example, auto mechanic, high school teacher, homemaker, laborer, lathe operator, shoe salesman, army sergeant.)

FATHER'S TYPE OF WORK:

metal
NIM Factory worker

MOTHER'S TYPE OF WORK

TODAY'S DATE

YOUR BIRTHDATE

Mo. Date Yr.

Mo. 39 Date 43 Yr. 86

GRADE IN SCHOOL

IF YOU ARE WORKING, PLEASE STATE YOUR TYPE OF WORK:

NOT ATTENDING SCHOOL ☐

Please fill out this form to reflect your views, even if other people might not agree. Feel free to print additional comments beside each item and in the spaces provided on pages 2 and 4.

I. Please list the sports you most like to take part in. For example: swimming, baseball, skating, skate boarding, bike riding, fishing, etc.

☐ None

a. football
b. Tennis
c. cricket

Compared to others of your age, about how much time do you spend in each?

Less Than Average Average More Than Average

☒ ☐ ☐
☐ ☒ ☐
☐ ☒ ☐

Compared to others of your age, how well do you do each one?

Below Average Average Above Average

☐ ☒ ☐
☐ ☐ ☒
☒ ☐ ☐

II. Please list your favorite hobbies, activities, and games, other than sports. For example: cards, books, piano, cars, crafts, etc. (Do not include listening to radio or TV.)

☐ None

a. Junior monorly
b. wild work
c. cartoon planet

Compared to others of your age, about how much time do you spend in each?

Less Than Average Average More Than Average

☐ ☒ ☐
☒ ☐ ☐
☐ ☐ ☒

Compared to others of your age, how well do you do each one?

Below Average Average Above Average

☐ ☐ ☐
☐ ☒ ☒
☒ ☐ ☐

III. Please list any organizations, clubs, teams or groups you belong to.

☐ None

a. out of school
b. go to a football club
c.

Compared to others of your age, how active are you in each?

Less Active Average More Active

☐ ☐ ☐
☐ ☐ ☐
☐ ☐ ☐

IV. Please list any jobs or chores you have. For example: paper route, babysitting, making bed, working in store, etc. (Include both paid and unpaid jobs and chores.)

☐ None

a. doing the washing up
b.
c.

Compared to others of your age, how well do you carry them out?

Below Average Average Above Average

☐ ☒ ☐
☐ ☐ ☐
☐ ☐ ☐

Please Print

1. About how many close friends do you have? ☐ None ☐ 1 ☐ 2 or 3 ☒ 4 or more
(Do not include brothers & sisters)

2. About how many times a week do you do things with any friends outside of regular school hours?
(Do not include brothers & sisters) ☐ less than 1 ☒ 1 or 2 ☐ 3 or more

Compared to others of your age, how well do you:

	Worse	About Average	Better	
a. Get along with your brothers & sisters?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/> I have no brothers or sisters
b. Get along with other kids?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
c. Get along with your parents?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	
d. Do things by yourself?	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	

Performance in academic subjects. ☒ I do not attend school because _____

Check a box for each subject that you take

	Failing	Below Average	Average	Above Average
a. English or Language Arts	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
b. History or Social Studies	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
c. Arithmetic or Math	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
d. Science	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
e. <u>computer</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
f. <u>gym</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
g. <u>shop</u>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Other academic subjects—for example: computer courses, foreign language, business. Do not include gym, shop, driver's ed., etc.

Do you have any illness, disability, or handicap? ☒ No ☐ Yes—please describe:

Please describe any concerns or problems you have about school:

None

Please describe any other concerns you have:

None

Please describe the best things about yourself:

~~None~~ I wear a school uniform and I do hard work.

Below is a list of items that describe kids. For each item that describes you **now or within the past 6 months**, please circle the **2** if the item is **very true or often true** of you. Circle the **1** if the item is **somewhat or sometimes true** of you. If the item is **not true** of you, circle the **0**.

Please Print

0 = Not True

1 = Somewhat or Sometimes True

2 = Very True or Often True

- | | | | |
|---|---|---|--|
| 0 | 1 | 2 | 1. I act too young for my age |
| 0 | 1 | 2 | 2. I have an allergy (describe): _____ |
| 0 | 1 | 2 | 3. I argue a lot |
| 0 | 1 | 2 | 4. I have asthma |
| 0 | 1 | 2 | 5. I act like the opposite sex |
| 0 | 1 | 2 | 6. I like animals |
| 0 | 1 | 2 | 7. I brag |
| 0 | 1 | 2 | 8. I have trouble concentrating or paying attention |
| 0 | 1 | 2 | 9. I can't get my mind off certain thoughts (describe): _____ |
| 0 | 1 | 2 | 10. I have trouble sitting still |
| 0 | 1 | 2 | 11. I'm too dependent on adults |
| 0 | 1 | 2 | 12. I feel lonely |
| 0 | 1 | 2 | 13. I feel confused or in a fog |
| 0 | 1 | 2 | 14. I cry a lot |
| 0 | 1 | 2 | 15. I am pretty honest |
| 0 | 1 | 2 | 16. I am mean to others |
| 0 | 1 | 2 | 17. I daydream a lot |
| 0 | 1 | 2 | 18. I deliberately try to hurt or kill myself |
| 0 | 1 | 2 | 19. I try to get a lot of attention |
| 0 | 1 | 2 | 20. I destroy my own things |
| 0 | 1 | 2 | 21. I destroy things belonging to others |
| 0 | 1 | 2 | 22. I disobey my parents |
| 0 | 1 | 2 | 23. I disobey at school |
| 0 | 1 | 2 | 24. I don't eat as well as I should |
| 0 | 1 | 2 | 25. I don't get along with other kids |
| 0 | 1 | 2 | 26. I don't feel guilty after doing something I shouldn't |
| 0 | 1 | 2 | 27. I am jealous of others |
| 0 | 1 | 2 | 28. I am willing to help others when they need help |
| 0 | 1 | 2 | 29. I am afraid of certain animals, situations, or places, other than school (describe): _____ |
| 0 | 1 | 2 | 30. I am afraid of going to school |
| 0 | 1 | 2 | 31. I am afraid I might think or do something bad |
| 0 | 1 | 2 | 32. I feel that I have to be perfect |
| 0 | 1 | 2 | 33. I feel that no one loves me |
| 0 | 1 | 2 | 34. I feel that others are out to get me |
| 0 | 1 | 2 | 35. I feel worthless or inferior |
| 0 | 1 | 2 | 36. I accidentally get hurt a lot |
| 0 | 1 | 2 | 37. I get in many fights |
| 0 | 1 | 2 | 38. I get teased a lot |
| 0 | 1 | 2 | 39. I hang around with kids who get in trouble |

- | | | | |
|---|---|---|--|
| 0 | 1 | 2 | 40. I hear sounds or voices that other people think aren't there (describe): _____ |
| 0 | 1 | 2 | 41. I act without stopping to think |
| 0 | 1 | 2 | 42. I would rather be alone than with others |
| 0 | 1 | 2 | 43. I lie or cheat |
| 0 | 1 | 2 | 44. I bite my fingernails |
| 0 | 1 | 2 | 45. I am nervous or tense |
| 0 | 1 | 2 | 46. Parts of my body twitch or make nervous movements (describe): _____ |
| 0 | 1 | 2 | 47. I have nightmares |
| 0 | 1 | 2 | 48. I am not liked by other kids |
| 0 | 1 | 2 | 49. I can do certain things better than most kids |
| 0 | 1 | 2 | 50. I am too fearful or anxious |
| 0 | 1 | 2 | 51. I feel dizzy |
| 0 | 1 | 2 | 52. I feel too guilty |
| 0 | 1 | 2 | 53. I eat too much |
| 0 | 1 | 2 | 54. I feel overtired |
| 0 | 1 | 2 | 55. I am overweight |
| 0 | 1 | 2 | 56. Physical problems without known medical cause: |
| 0 | 1 | 2 | a. Aches or pains (not stomach or headaches) |
| 0 | 1 | 2 | b. Headaches |
| 0 | 1 | 2 | c. Nausea, feel sick |
| 0 | 1 | 2 | d. Problems with eyes (not if corrected by glasses) (describe): _____ |
| 0 | 1 | 2 | e. Rashes or other skin problems |
| 0 | 1 | 2 | f. Stomachaches or cramps |
| 0 | 1 | 2 | g. Vomiting, throwing up |
| 0 | 1 | 2 | h. Other (describe): _____ |
| 0 | 1 | 2 | 57. I physically attack people |
| 0 | 1 | 2 | 58. I pick my skin or other parts of my body (describe): _____ |
| 0 | 1 | 2 | 59. I can be pretty friendly |
| 0 | 1 | 2 | 60. I like to try new things |
| 0 | 1 | 2 | 61. My school work is poor |
| 0 | 1 | 2 | 62. I am poorly coordinated or clumsy |
| 0 | 1 | 2 | 63. I would rather be with older kids than with kids my own age |

0 = Not True

1 = Somewhat or Sometimes True

2 = Very True or Often True

- 1 2 64. I would rather be with younger kids than with kids my own age
- 1 2 65. I refuse to talk
- 1 2 66. I repeat certain acts over and over (describe): _____

- 1 2 67. I run away from home
- 1 2 68. I scream a lot
- 1 2 69. I am secretive or keep things to myself
- 1 2 70. I see things that other people think aren't there (describe): _____

- 1 2 71. I am self-conscious or easily embarrassed
- 1 2 72. I set fires
- 1 2 73. I can work well with my hands
- 1 2 74. I show off or clown
- 1 2 75. I am shy
- 1 2 76. I sleep less than most kids
- 1 2 77. I sleep more than most kids during day and/or night (describe): _____

- 1 2 78. I have a good imagination
- 1 2 79. I have a speech problem (describe): _____

- 1 2 80. I stand up for my rights
- 1 2 81. I steal at home
- 1 2 82. I steal from places other than home
- 1 2 83. I store up things I don't need (describe): _____

- 1 2 84. I do things other people think are strange (describe): _____

- 0 1 2 85. I have thoughts that other people would think are strange (describe): _____

- 0 1 2 86. I am stubborn
- 0 1 2 87. My moods or feelings change suddenly
- 0 1 2 88. I enjoy being with other people
- 0 1 2 89. I am suspicious
- 0 1 2 90. I swear or use dirty language
- 0 1 2 91. I think about killing myself
- 0 1 2 92. I like to make others laugh
- 0 1 2 93. I talk too much
- 0 1 2 94. I tease others a lot
- 0 1 2 95. I have a hot temper
- 0 1 2 96. I think about sex too much
- 0 1 2 97. I threaten to hurt people
- 0 1 2 98. I like to help others
- 0 1 2 99. I am too concerned about being neat or clean
- 0 1 2 100. I have trouble sleeping (describe): _____

- 0 1 2 101. I cut classes or skip school
- 0 1 2 102. I don't have much energy
- 0 1 2 103. I am unhappy, sad, or depressed
- 0 1 2 104. I am louder than other kids
- 0 1 2 105. I use alcohol or drugs for nonmedical purposes (describe): _____

- 0 1 2 106. I try to be fair to others
- 0 1 2 107. I enjoy a good joke
- 0 1 2 108. I like to take life easy
- 0 1 2 109. I try to help other people when I can
- 0 1 2 110. I wish I were of the opposite sex
- 0 1 2 111. I keep from getting involved with others
- 0 1 2 112. I worry a lot

Please write down anything else that describes your feelings, behavior, or interests

TEACHER'S REPORT FORM FOR AGES 5-18

For office use only
ID #

Please Print

Your answers will be used to compare the pupil with other pupils whose teachers have completed similar forms. The information from this form will also be used for comparison with other information about this pupil. Please answer as well as you can, even if you lack full information. Scores on individual items will be combined to identify general patterns of behavior. Feel free to print additional comments beside each item and in the spaces provided on page 2.

PUPIL'S FULL NAME		FIRST	MIDDLE	LAST	PARENTS' USUAL TYPE OF WORK, even if not working now (Please be as specific as you can—for example, auto mechanic, high school teacher, homemaker, laborer, lathe operator, shoe salesman, army sergeant.) FATHER'S TYPE OF WORK: _____ MOTHER'S TYPE OF WORK: _____ THIS FORM FILLED OUT BY: <input type="checkbox"/> Teacher (full name) _____ <input type="checkbox"/> Counselor (full name) _____ <input type="checkbox"/> Other (specify position & give full name): _____
PUPIL'S SEX		PUPIL'S AGE		ETHNIC GROUP OR RACE	
<input type="checkbox"/> Boy <input type="checkbox"/> Girl					
TODAY'S DATE		PUPIL'S BIRTHDATE (if known)			
Mo. _____ Date _____ Yr. _____		Mo. _____ Date _____ Yr. _____			
GRADE IN SCHOOL	NAME AND ADDRESS OF SCHOOL				

I. For how many months have you known this pupil? _____ months

II. How well do you know him/her? 1. ☐ Not Well 2. ☐ Moderately Well 3. ☐ Very Well

III. How much time does he/she spend in your class or service per week?

IV. What kind of class or service is it? (Please be specific, e.g., regular 5th grade, 7th grade math, learning disabled, counseling, etc.)

V. Has he/she ever been referred for special class placement, services, or tutoring?

☐ Don't Know 0. ☐ No 1. ☐ Yes—what kind and when?

VI. Has he/she repeated any grades?

☐ Don't Know 0. ☐ No 1. ☐ Yes—grades and reasons

VII. **Current school performance**—list academic subjects and check box that indicates pupil's performance for each subject:

Academic subject	1. Far below grade	2. Somewhat below grade	3. At grade level	4. Somewhat above grade	5. Far above grade
1. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

VIII. Compared to typical pupils of the same age:	1. Much less	2. Somewhat less	3. Slightly less	4. About average	5. Slightly more	6. Somewhat more	7. Much more
1. How hard is he/she working?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. How appropriately is he/she behaving?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. How much is he/she learning?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. How happy is he/she?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

IX. Most recent achievement test scores (optional).

Name of test	Subject	Date	Percentile or grade level obtained

X. IQ, readiness, or aptitude tests (optional).

Name of test	Date	IQ or equivalent scores

Does this pupil have any illness or disability (either physical or mental)? ☐ No ☐ Yes—please describe:

What concerns you most about this pupil?

Please describe the best things about this pupil:

Please feel free to write any comments about this pupil's work, behavior, or potential, using extra pages if necessary.

Please Print

Below is a list of items that describe pupils. For each item that describes the pupil **now or within the past 2 months**, please circle the **2** if the item is **very true or often true** of the pupil. Circle the **1** if the item is **somewhat or sometimes true** of the pupil. If the item is **not true** of the pupil, circle the **0**. Please answer all items as well as you can, even if some do not seem to apply to this pupil.

0 = Not True (as far as you know)

1 = Somewhat or Sometimes True

2 = Very True or Often True

- | | | | |
|---|---|---|--|
| 0 | 1 | 2 | 1. Acts too young for his/her age |
| 0 | 1 | 2 | 2. Hums or makes other odd noises in class |
| 0 | 1 | 2 | 3. Argues a lot |
| 0 | 1 | 2 | 4. Fails to finish things he/she starts |
| 0 | 1 | 2 | 5. Behaves like opposite sex |
| 0 | 1 | 2 | 6. Defiant, talks back to staff |
| 0 | 1 | 2 | 7. Bragging, boasting |
| 0 | 1 | 2 | 8. Can't concentrate, can't pay attention for long |
| 0 | 1 | 2 | 9. Can't get his/her mind off certain thoughts; obsessions (describe): _____ |
| 0 | 1 | 2 | 10. Can't sit still, restless, or hyperactive |
| 0 | 1 | 2 | 11. Clings to adults or too dependent |
| 0 | 1 | 2 | 12. Complains of loneliness |
| 0 | 1 | 2 | 13. Confused or seems to be in a fog |
| 0 | 1 | 2 | 14. Cries a lot |
| 0 | 1 | 2 | 15. Fidgets |
| 0 | 1 | 2 | 16. Cruelty, bullying, or meanness to others |
| 0 | 1 | 2 | 17. Daydreams or gets lost in his/her thoughts |
| 0 | 1 | 2 | 18. Deliberately harms self or attempts suicide |
| 0 | 1 | 2 | 19. Demands a lot of attention |
| 0 | 1 | 2 | 20. Destroys his/her own things |
| 0 | 1 | 2 | 21. Destroys property belonging to others |
| 0 | 1 | 2 | 22. Difficulty following directions |
| 0 | 1 | 2 | 23. Disobedient at school |
| 0 | 1 | 2 | 24. Disturbs other pupils |
| 0 | 1 | 2 | 25. Doesn't get along with other pupils |
| 0 | 1 | 2 | 26. Doesn't seem to feel guilty after misbehaving |
| 0 | 1 | 2 | 27. Easily jealous |
| 0 | 1 | 2 | 28. Eats or drinks things that are not food— don't include sweets (describe): _____ |
| 0 | 1 | 2 | 29. Fears certain animals, situations, or places other than school (describe): _____ |
| 0 | 1 | 2 | 30. Fears going to school |

- | | | | |
|---|---|---|---|
| 0 | 1 | 2 | 31. Fears he/she might think or do something bad |
| 0 | 1 | 2 | 32. Feels he/she has to be perfect |
| 0 | 1 | 2 | 33. Feels or complains that no one loves him/her |
| 0 | 1 | 2 | 34. Feels others are out to get him/her |
| 0 | 1 | 2 | 35. Feels worthless or inferior |
| 0 | 1 | 2 | 36. Gets hurt a lot, accident-prone |
| 0 | 1 | 2 | 37. Gets in many fights |
| 0 | 1 | 2 | 38. Gets teased a lot |
| 0 | 1 | 2 | 39. Hangs around with others who get in trouble |
| 0 | 1 | 2 | 40. Hears sounds or voices that aren't there (describe): _____ |
| 0 | 1 | 2 | 41. Impulsive or acts without thinking |
| 0 | 1 | 2 | 42. Would rather be alone than with others |
| 0 | 1 | 2 | 43. Lying or cheating |
| 0 | 1 | 2 | 44. Bites fingernails |
| 0 | 1 | 2 | 45. Nervous, high-strung, or tense |
| 0 | 1 | 2 | 46. Nervous movements or twitching (describe): _____ |
| 0 | 1 | 2 | 47. Overconforms to rules |
| 0 | 1 | 2 | 48. Not liked by other pupils |
| 0 | 1 | 2 | 49. Has difficulty learning |
| 0 | 1 | 2 | 50. Too fearful or anxious |
| 0 | 1 | 2 | 51. Feels dizzy |
| 0 | 1 | 2 | 52. Feels too guilty |
| 0 | 1 | 2 | 53. Talks out of turn |
| 0 | 1 | 2 | 54. Overtired |
| 0 | 1 | 2 | 55. Overweight |
| 0 | 1 | 2 | 56. Physical problems without known medical cause: |
| 0 | 1 | 2 | a. Aches or pains (not stomach or headaches) |
| 0 | 1 | 2 | b. Headaches |
| 0 | 1 | 2 | c. Nausea, feel sick |
| 0 | 1 | 2 | d. Problems with eyes (not if corrected by glasses) (describe): _____ |
| 0 | 1 | 2 | e. Rashes or other skin problems |
| 0 | 1 | 2 | f. Stomachaches or cramps |
| 0 | 1 | 2 | g. Vomiting, throwing up |
| 0 | 1 | 2 | h. Other (describe): _____ |

Please Print

0 = Not True (as far as you know)

1 = Somewhat or Sometimes True

2 = Very True or Often True

- | | | | |
|---|---|---|---|
| 0 | 1 | 2 | 57. Physically attacks people |
| 0 | 1 | 2 | 58. Picks nose, skin, or other parts of body
(describe): _____
_____ |
| | | | |
| 0 | 1 | 2 | 59. Sleeps in class |
| 0 | 1 | 2 | 60. Apathetic or unmotivated |
| | | | |
| 0 | 1 | 2 | 61. Poor school work |
| 0 | 1 | 2 | 62. Poorly coordinated or clumsy |
| | | | |
| 0 | 1 | 2 | 63. Prefers being with older children or youths |
| 0 | 1 | 2 | 64. Prefers being with younger children |
| | | | |
| 0 | 1 | 2 | 65. Refuses to talk |
| 0 | 1 | 2 | 66. Repeats certain acts over and over; compulsions
(describe): _____
_____ |
| | | | |
| 0 | 1 | 2 | 67. Disrupts class discipline |
| 0 | 1 | 2 | 68. Screams a lot |
| | | | |
| 0 | 1 | 2 | 69. Secretive, keeps things to self |
| 0 | 1 | 2 | 70. Sees things that aren't there (describe): _____
_____ |
| | | | |
| 0 | 1 | 2 | 71. Self-conscious or easily embarrassed |
| 0 | 1 | 2 | 72. Messy work |
| | | | |
| 0 | 1 | 2 | 73. Behaves irresponsibly (describe): _____
_____ |
| | | | |
| 0 | 1 | 2 | 74. Showing off or clowning |
| 0 | 1 | 2 | 75. Shy or timid |
| 0 | 1 | 2 | 76. Explosive and unpredictable behavior |
| | | | |
| 0 | 1 | 2 | 77. Demands must be met immediately, easily
frustrated |
| 0 | 1 | 2 | 78. Inattentive, easily distracted |
| | | | |
| 0 | 1 | 2 | 79. Speech problem (describe): _____
_____ |
| | | | |
| 0 | 1 | 2 | 80. Stares blankly |
| | | | |
| 0 | 1 | 2 | 81. Feels hurt when criticized |
| | | | |
| 0 | 1 | 2 | 82. Steals |
| 0 | 1 | 2 | 83. Stores up things he/she doesn't need (describe): _____

_____ |

- | | | | |
|---|---|---|---|
| 0 | 1 | 2 | 84. Strange behavior (describe): _____
_____ |
| | | | |
| 0 | 1 | 2 | 85. Strange ideas (describe): _____
_____ |
| | | | |
| 0 | 1 | 2 | 86. Stubborn, sullen, or irritable |
| | | | |
| 0 | 1 | 2 | 87. Sudden changes in mood or feelings |
| 0 | 1 | 2 | 88. Sulks a lot |
| | | | |
| 0 | 1 | 2 | 89. Suspicious |
| 0 | 1 | 2 | 90. Swearing or obscene language |
| | | | |
| 0 | 1 | 2 | 91. Talks about killing self |
| 0 | 1 | 2 | 92. Underachieving, not working up to potential |
| | | | |
| 0 | 1 | 2 | 93. Talks too much |
| 0 | 1 | 2 | 94. Teases a lot |
| | | | |
| 0 | 1 | 2 | 95. Temper tantrums or hot temper |
| 0 | 1 | 2 | 96. Seems preoccupied with sex |
| | | | |
| 0 | 1 | 2 | 97. Threatens people |
| 0 | 1 | 2 | 98. Tardy to school or class |
| | | | |
| 0 | 1 | 2 | 99. Too concerned with neatness or cleanliness |
| 0 | 1 | 2 | 100. Fails to carry out assigned tasks |
| | | | |
| 0 | 1 | 2 | 101. Truancy or unexplained absence |
| 0 | 1 | 2 | 102. Underactive, slow moving, or lacks energy |
| | | | |
| 0 | 1 | 2 | 103. Unhappy, sad, or depressed |
| 0 | 1 | 2 | 104. Unusually loud |
| | | | |
| 0 | 1 | 2 | 105. Uses alcohol or drugs for nonmedical purposes
(describe): _____
_____ |
| | | | |
| 0 | 1 | 2 | 106. Overly anxious to please |
| | | | |
| 0 | 1 | 2 | 107. Dislikes school |
| 0 | 1 | 2 | 108. Is afraid of making mistakes |
| | | | |
| 0 | 1 | 2 | 109. Whining |
| 0 | 1 | 2 | 110. Unclean personal appearance |
| | | | |
| 0 | 1 | 2 | 111. Withdrawn, doesn't get involved with others |
| 0 | 1 | 2 | 112. Worries |
| | | | |
| | | | 113. Please write in any problems the pupil has that
were not listed above:

_____ |
| | | | |
| 0 | 1 | 2 | _____ |
| | | | |
| 0 | 1 | 2 | _____ |
| | | | |
| 0 | 1 | 2 | _____ |

MSCS

Multidimensional Self Concept Scale

Bruce A. Bracken

RECORD BOOKLET

Section I. Identifying Information

Time/ID No.

Address

Parents' Name

School/Agency

Referred by

Place of testing

Tested by

Place

B

W

Other

Spanish Origin

Yes

No

Year

Month

Day

Date Tested

/

/

Age

Date of Birth

/

/

Sex

Chronological Age

/

/

Grade

Section II. Directions and Scales

Please rate the following statements according to how well the statement applies to you. There are no right or wrong answers, but it is important that you rate each statement according to how you honestly feel. Be sure to be honest with yourself as you consider the statement you are rating. To mark your answer, simply circle the letters that correspond with your feelings toward the statement. Each statement should be rated as:

Strongly Agree
(SA)

Agree
(A)

Disagree
(D)

Strongly Disagree
(SD)

S SCALE

Strongly Agree
(SA)

Agree
(A)

Disagree
(D)

Strongly Disagree
(SD)

SCORE

1. I am usually a lot of fun to be with

SA A D SD

2. People do not seem interested in talking with me

SA - A D SD

3. I am too shy

SA A D SD

4. Most people like me

SA A D SD

5. People avoid me

SA A D SD

6. A lot of people make fun of me

SA A D SD

7. I am not accepted by people who know me

SA A D SD

8. Most people think I am interesting

SA A D SD

9. People enjoy being with me

SA A D SD

10. Most of the time I feel ignored

SA A D SD

11. I feel desired by members of the opposite sex

SA A D SD

12. No one seems to laugh at my jokes

SA A D SD

13. Most people appreciate me just the way I am

SA A D SD

14. I often feel like I am left out of things

SA A D SD

15. People tell lies about me

SA A D SD

16. I have a lot of friends

SA A D SD

17. I spend a lot of time feeling lonely

SA A D SD

18. I am never sure how to act when I am with people I don't know well

SA A D SD

19. People tell me their secrets

SA A D SD

20. People pick on me

SA A D SD

21. People do not seem to notice me

SA A D SD

22. I get a lot of phone calls from friends

SA A D SD

23. Many people have a low opinion of me

SA A D SD

24. I let people bully me too much

SA A D SD

25. People have to get to know me before they like me

SA A D SD

S Scale Total Raw Score

C SCALE

Strongly Agree
(SA)

Agree
(A)

Disagree
(D)

Strongly Disagree
(SD)

SCORE

26. I am honest

SA A D SD

27. Too often I say the wrong thing

SA - A D SD

28. I am too lazy

SA A D SD

29. I have a good sense of humor

SA A D SD

30. I am basically a weak person

SA A D SD

31. I feel that most people respect me

SA A D SD

32. I am not very good at speaking my mind

SA A D SD

33. I am assertive when I need to be

SA A D SD

34. I am unlucky

SA A D SD

35. I am very self confident

SA A D SD

36. I don't seem to have any control over my life

SA A D SD

37. I frequently put off doing important things until it is too late

SA A D SD

38. I give people good reason to trust me

SA A D SD

39. I am not as good as I should be

SA A D SD

40. I don't keep quiet when I should

SA A D SD

41. I am successful at most things

SA A D SD

42. I handle my personal business responsibly

SA A D SD

43. I lack common sense

SA A D SD

44. I always seem to be in trouble

SA A D SD

45. I can do most things pretty well

SA A D SD

46. I am not very smart

SA A D SD

47. I am a coward in many ways

SA A D SD

48. Others believe that I will make something of myself

SA A D SD

49. Too often I do dumb things without thinking

SA A D SD

50. I waste money foolishly

SA A D SD

C Scale Total Raw Score

AFF SCALE

Strongly Agree
(SA)

Agree
(A)

Disagree
(D)

Strongly Disagree
(SD)

SCORE

51. I enjoy life

SA A D SD

52. I am afraid of many things

SA A D SD

53. There are many things I would like to change about myself

SA A D SD

54. I am not able to laugh at myself very easily

SA A D SD

55. I am not a happy person

SA A D SD

56. I am proud of myself

SA A D SD

57. I feel like a failure

SA A D SD

58. My life is discouraging

SA A D SD

59. I am happy with myself just the way I am

SA A D SD

60. I am too emotional

SA A D SD

61. I have good self control

SA A D SD

62. I often disappoint myself

SA A D SD

63. My life is unstable

SA A D SD

64. I have a positive outlook on life

SA A D SD

65. I am frequently confused about my feelings

SA A D SD

66. Sometimes I feel worthless

SA A D SD

67. I often feel ashamed of things I have done

SA A D SD

68. I frequently feel helpless

SA A D SD

69. I feel loved

SA A D SD

70. I wish I could be someone else

SA A D SD

71. I feel insecure

SA A D SD

72. I am a good person

SA A D SD

73. I am not as happy as I appear

SA A D SD

74. I am usually very relaxed

SA A D SD

75. There are times when I don't like myself

SA A D SD

AFF Scale Total Raw Score

AC SCALE

AC SCALE					
Strongly Agree (SA)	Agree (A)	Disagree (D)	Strongly Disagree (SD)		
				SCORE	
76. Classmates usually like my ideas	SA	A	D	SD	
77. I frequently feel unprepared for class	SA	A	D	SD	
78. I am good at mathematics	SA	A	D	SD	
79. Learning is difficult for me	SA	A	D	SD	
80. I usually do well on tests	SA	A	D	SD	
81. I am proud of my school work	SA	A	D	SD	
82. I can spell better than most people my age	SA	A	D	SD	
83. I read as well as most people my age	SA	A	D	SD	
84. I don't think very quickly	SA	A	D	SD	
85. I work harder than most of my classmates	SA	A	D	SD	
86. I don't understand much of what I read	SA	A	D	SD	
87. I learn fairly easily	SA	A	D	SD	
88. I never seem to have good ideas	SA	A	D	SD	
89. My teachers like my classroom behavior	SA	A	D	SD	
90. I often feel dumb	SA	A	D	SD	
91. Most of my teachers seem to like me	SA	A	D	SD	
92. I have poor study habits	SA	A	D	SD	
93. Science is easy for me	SA	A	D	SD	
94. I am uncomfortable in school	SA	A	D	SD	
95. I usually work very hard	SA	A	D	SD	
96. Most people would rather work with me than someone else	SA	A	D	SD	
97. My teachers have a low opinion of me	SA	A	D	SD	
98. Most subjects are pretty easy for me	SA	A	D	SD	
99. I am not very creative	SA	A	D	SD	
100. I usually feel good about my written work	SA	A	D	SD	
AC Scale Total Raw Score					

F SCALE

Strongly Agree
(SA)

Agree
(A)

Disagree
(D)

Strongly Disagree
(SD)

SCORE

101. My parents care about my happiness

SA A D SD

102. My family makes me feel loved

SA A D SD

103. My family ruins everything for me

SA A D SD

104. In my family, we take care of each other

SA A D SD

105. I feel appreciated by my family

SA A D SD

106. I have fun with my family

SA A D SD

107. I wish I could trade families with someone else

SA A D SD

108. My parents are interested in me

SA A D SD

109. My parents don't trust me

SA A D SD

110. My home is warm and caring

SA A D SD

111. My parents do not like my being around them

SA A D SD

112. My parents help me when I need it

SA A D SD

113. I am an important member of my family

SA A D SD

114. My parents are proud of me

SA A D SD

115. My family is no good

SA A D SD

116. Nothing I do seems to please my parents

SA A D SD

117. My parents attend events that are important to me

SA A D SD

118. My parents believe in me

SA A D SD

119. I am proud of my family

SA A D SD

120. My parents care about my education

SA A D SD

121. My family is one of the most important parts of my life

SA A D SD

122. My parents love me just as I am

SA A D SD

123. I don't know why my family stays together

SA A D SD

124. My parents care about my future

SA A D SD

125. My home is not a happy place

SA A D SD

F Scale Total Raw Score

P-SCALE

Strongly Agree
(SA)

Agree
(A)

Disagree
(D)

Strongly Disagree
(SD)

SCORE

126. I feel good

SA A D SD

127. I am attractive

SA A D SD

128. I am in poor shape

SA A D SD

129. When I look in the mirror, I like what I see

SA A D SD

130. I tire too quickly

SA A D SD

131. I have nice looking teeth

SA A D SD

132. I look nice in just about anything I wear

SA A D SD

133. I am ugly

SA A D SD

134. I am stronger than most people

SA A D SD

135. I have a nice figure

SA A D SD

136. I am healthy

SA A D SD

137. I feel good about how I look

SA A D SD

138. I am good at most sports

SA A D SD

139. I do not like how my clothes fit me

SA A D SD

140. I am typically chosen among the last for team sports

SA A D SD

141. I am physically fit

SA A D SD

142. My hair never seems to look very good

SA A D SD

143. My skin is attractive

SA A D SD

144. I do not like to be seen in a swimsuit

SA A D SD

145. There are parts of my body that I try to keep others from noticing

SA A D SD

146. My clothes look good on me

SA A D SD

147. I do not seem to have the energy to do very much

SA A D SD

148. My weight is just about where it should be

SA A D SD

149. I would change my looks if I could

SA A D SD

150. I am graceful

SA A D SD

P Scale Total Raw Score

DO NOT WRITE ON THIS PAGE

Section III Norm-Referenced Interpretation					Section IV Ipsative Interpretation		
Raw Score	Standard Score	Confidence Interval	Classification	%ile Rank	Standard Score	Difference Score	.05/.01 Classification

Section V Profile of Scale Scores							
Standard Score	Social	Competence	Affect	Academic	Family	Physical	Total Test
145							
—							
—							
130							
—							
—							
115							
—							
—							
100							
—							
—							
85							
—							
—							
70							
—							
—							
55							

Appendix 3: Direct Observation Protocol

Observer Guidelines

- 16 children have been selected from each class to give a range of behaviours.
- These children will be observed for 8 hours in total, 4 as the research starts and 4 after it has finished.
- The 4 hours are split, so that you will observe 2 morning sessions on different days and 2 afternoon sessions, - the afternoon session may follow a morning session on the same day

The hour session is divided into four 15 minute periods.

In each period the children's behaviours are recorded on 3 measures

- a) On Task Observation
- b) Disruption Tally
- c) Running Record

The observation period takes this pattern:-

- a) - 2 minutes
- b) - 6 minutes
- a) - 2 minutes
- c) - 5 minutes

Repeated four times over the session.

The session is concluded with a teacher de-briefing questionnaire - this should take 5 - 10 minutes.

- The more information the better - put lots down even if feel a little unfair. It will come out in the wash.
- Consistency - no 'going gooey'.
- The first response is the best one.
- The emphasis is on the instantaneous nature of observation, what each child is doing at that particular moment.

GUIDE TO THE 3 MEASURES USED

A. On Task Observation

There are five categories of behaviour to register:-

V = On Task

X = Off Task

D = Disruptive

W = Waiting

O = Other (mark prosocial behaviour as , withdrawn as N)

Here's what you do:-

1. Check that you know where each child on the list is sitting.
2. Note the time.
3. Count to 4
4. Locate child no. 1 and register his behaviour on the schedule
3. Count to 4
4. Locate child No. 2 and register his behaviour etc.

It should take 2 minutes to register 16 children.

Emphasis on the instantaneous nature of the observation.

In the space for notes make a brief observation about children observed to be lost/withdrawn/depressed. Ask for teacher's comments on any children fitting this category

CLASSROOM OBSERVATION

Aside from the extremely tedious forms I am sure you are, or will be, sick of, the main data used in evaluation the research is gathered through classroom observation.

This will involve a suitably experienced research assistant (not a teacher) coming into your class for 4 periods at the beginning of term and 4 periods at the end of the intervention.

It is this person's job to sit quietly in class and complete a schedule that classifies the behaviour of children in your class. The teacher's input, style and effectiveness are not being assessed through this observation. What is essential information is the way the children react within a group and how they function as part of a class.

I will be happy to supply copies of the schedule if you are interested, but I have included the teacher de-briefing form that is the concluding part of the observation for you to look at, because it is the part that relates directly to you. If there are any modifications that you would make, please let me know.

The ideal situation would be for the observation to take place in a both morning and afternoon session before and after the intervention. Please plan the lesson as you would do normally, because the more information that can be recorded the better, it is not a perfect 10 for classroom control that will yield the most useful information. The observer will need a seating plan for that session with the names of the children written on.

If there are any difficulties at all about the notion of classroom observation, then please let me know and we can talk about it. Many thanks, as ever, to you for the commitment above and beyond the call of duty. It is greatly appreciated.

OBSERVATION PROCEDURE

Contact Head

- Go ahead

Contact Teacher

- agree days and times
- get seating plans for those lessons
- principles: this is an ordinary lesson with usual mix of teacher input, individual and group work
- it is not the teacher that is being assessed but the children that are being observed
- Choose lesson when less likely to be on carpet for whole class period

MASTER SHEET FOR CLASSROOM OBSERVATION IN
SCHOOL ON

Attached to this sheet should be 4X Observation sheet A & C and 4X Observation sheet B

LESSON CONTENT
Mark content and any changes in organisation/subject matter on the time line below with an indication of the time

Start	End
<div></div>	

V = ON TASK
X = OFF TASK

D = DISRUPTIVE
W = WAITING

U = PRO-SOCIAL BEHAVIOUR.

SCHEDULE FOR CLASSROOM OBSERVATION IN _____
SCHOOL ON _____ PERIOD NO. _____

(Fill in names and copy 4 times for each session)

A. On Task Observation

[illegible]

A. Notes

C. Running Record

B Help - Helping another pupil
Toot - Talking out of turn

HOC - Hindering others.
OOS - Out of Seat

AV/P - Aggression - Verbal or Physical (Mark V or P)

[illegible][illegible][illegible][illegible][illegible]

NAME	H	T	H	O	A	NAME	H	T	H	O	A
E	O	O	O	V		E	O	O	O	V	
L	O	C	S	/		L	O	C	S	/	
P	T			A	P	P	T			A	P

