

**THE IMPACT OF PRIMARY SEAL SMALL GROUP
INTERVENTIONS (SILVER SET MATERIALS) ON
SOCIAL AND EMOTIONAL OUTCOMES FOR PUPILS**

Clare Otter

Thesis submitted to the University of Nottingham
for the degree of Doctor of Applied Educational
Psychology

November 2010

ABSTRACT

This study aimed to contribute to the small evidence-base on the Social and Emotional Aspects of Learning (SEAL) programme. SEAL is a school-based approach to developing children's social and emotional skills. It was launched by the government in 2005 and has been adopted by schools across the UK. This study focused on the small group aspect of SEAL, which is aimed at children who are targeted for additional practice with their social and emotional skills. The researcher used a non-equivalent control group quasi-experimental design to evaluate the effectiveness of two of the small group SEAL interventions, New Beginnings and Getting On and Falling Out. Pupils, parents and teachers completed questionnaires before and after the interventions and, in the case of New Beginnings, around six weeks after the intervention ended. The level of fidelity to the government guidance was assessed through observations and interviews. No positive results were found for the New Beginnings intervention group in comparison with the control group, but there was some support for the Getting On and Falling Out intervention; with improvements in teacher-rated empathy, total emotional literacy and pro-social behaviour. In common with previous research, no effects were found for parent-ratings or for children who had been selected to take part in the interventions as role models. The results are discussed in terms of implications for practitioner educational psychologists and suggestions are made for further studies in this under-researched area.

ACKNOWLEDGEMENTS

I am indebted to pupils, school staff and parents for their agreement to participate in the study. It would not have been possible without them.

I am also extremely grateful for the guidance and support with this project that was provided by staff at the University of Nottingham; particularly comments and reassurance from Dr Nathan Lambert and also discussions with Dr Nick Durbin and Dr Sarah Atkinson.

I would like to thank my colleagues in the 'PAST' Team at Central Bedfordshire Council for their understanding and support during my employment as a Trainee Educational Psychologist, particularly my supervisor Donna Wiggett.

Finally, special thanks go to my partner, Jun Shaw, and my friends and family for their unwavering moral support during this endeavour. Particular mention must go to my good friends in 'TEP '07' whose support over the three years has been crucial. I hope we continue to keep in touch throughout our careers.

I dedicate this thesis to my parents, Vic and Brenda Otter. I am so thankful for their unconditional love and support for my own social and emotional needs as a child. This still continues and I have come to realise that not every child is as lucky as I was in this respect. I look forward to giving their needs the priority in my life that they should have been given over recent months.

LIST OF CONTENTS

GLOSSARY	12
CHAPTER 1: INTRODUCTION	13
1.1 The context of and rationale for the research	13
1.1.1 <i>The initial training of educational psychologists</i>	13
1.1.2 <i>Evidence-based practice and the focus on children's outcomes</i>	14
1.1.3 <i>Meeting children's emotional needs in schools</i>	15
1.1.4 <i>Personal interest</i>	16
1.2 Aim of the study and overview of the chapters	17
CHAPTER 2: LITERATURE REVIEW	18
2.1 Introduction to the literature review chapter	18
2.2 The historical and theoretical underpinnings of Primary SEAL	19
2.2.1 <i>Overview of the Primary SEAL resource</i>	19
2.2.2 <i>Primary SEAL: aims and origins</i>	20
2.2.3 <i>Primary SEAL: theoretical underpinnings</i>	21
2.3 Psychological theories that influenced the content of SEAL	23
2.3.1 <i>Emotional intelligence (EI)</i>	23
<i>Origins and definitions</i>	23
<i>Popularisation</i>	24
<i>Controversy and fragmentation</i>	25
<i>The future?</i>	27
2.3.2 <i>Emotional literacy</i>	27
2.3.3 <i>Empathy</i>	29
2.3.4 <i>Social problem solving</i>	31
2.3.5 <i>Anger management</i>	34
2.3.6 <i>Cognitive-behavioural theories</i>	36
2.3.7 <i>Summary</i>	40
2.4 Educational practice that influenced the development of SEAL	40
2.4.1 <i>Social and emotional learning: definition and origins</i>	40

2.4.2 Reviews of SEL research that influenced the development of SEAL	41
2.5 The ‘waves of intervention’ model of meeting additional needs	44
2.6 Systematic reviews of research into SEL interventions	47
2.6.1 Details of the systematic search strategies	47
2.6.2 Evaluations of SEAL	52
Evaluations of wave two Primary SEAL	52
Evaluations of wave one Primary SEAL	56
Evaluations of Secondary SEAL	59
2.6.3 Evaluations of other wave two social and emotional interventions	60
2.6.4 Evaluations of other wave one social and emotional interventions	65
2.7 Evaluations of SEAL in Bedfordshire	68
2.8 Gaps in the literature and rationale for the present study	71
2.9 Research questions	75
 CHAPTER 3: METHODOLOGY	 76
3.1 Introduction to the methodology chapter	76
3.2 Epistemological issues	76
3.2.1 Positivism	76
3.2.2 Constructivism	77
3.2.3 Epistemological debate and the link with methodology	77
3.2.4 A pragmatic view	79
3.2.5 Real world research	80
3.2.6 Chosen stance	80
3.3 Research design	81
3.3.1 Research designs relating to the type of research questions posed	81
Experiments	82
Quasi-experiments	84
Pre-experimental designs	85
3.3.2 Issues of validity in experimental designs	86
Internal validity	86
External validity	87

3.3.3 Chosen research design	89
3.4 Data collection	91
3.4.1 Problems of reliability and validity in psychological research	91
Reliability	91
Validity	91
3.4.2 Issues in measuring social and emotional competencies	92
3.4.3 Chosen outcome measures	94
Strengths and Difficulties Questionnaire	95
Emotional Literacy Assessment Instrument	97
B/G Steem	99
3.5 Procedure	100
3.5.1 Timeline	100
3.5.2 Participants	100
Recruitment of participants	100
Schools	101
Pupils	105
Description of participants	106
Schools	106
Group leaders	110
Pupils	110
Response rates and attrition	112
3.5.3 Intervention	114
Materials	114
Frequency and duration	114
Location	114
Intervention integrity	115
Comparison group intervention	115
3.5.4 Ethical considerations	116
Consent	116
Deception	117
Debriefing	117

<i>Withdrawal</i>	117
<i>Confidentiality</i>	117
<i>Protection of participants</i>	118
<i>Observational research</i>	118
<i>Giving advice</i>	118
<i>Colleagues</i>	118
3.6 Data analysis	119
3.6.1 <i>Descriptive</i>	119
3.6.2 <i>Significance testing</i>	119
3.6.3 <i>Effect size</i>	121
 CHAPTER 4: RESULTS	 122
4.1 Introduction to the results chapter	122
4.2 Descriptive statistics	124
4.2.1 <i>Distribution of the data</i>	124
<i>New Beginnings</i>	125
<i>Getting On and Falling Out</i>	125
<i>Overall summary and implications</i>	126
4.2.2 <i>Measures of central tendency and variability</i>	126
4.3 Inferential statistics	134
4.3.1 <i>Between-group comparisons</i>	134
<i>Similarity of the groups at pre-test</i>	134
<i>Similarity of the groups at post-test</i>	136
<i>Summary of between-group comparisons</i>	137
4.3.2 <i>Pre to post-test change in scores</i>	140
<i>New Beginnings</i>	140
<i>Getting On and Falling Out</i>	143
<i>Summary of within-group comparisons</i>	145
4.3.3 <i>Analysis of follow-up measures</i>	145
4.3.4 <i>Analysis of sub-groups</i>	147
<i>New Beginnings</i>	148

<i>Getting On and Falling Out</i>	149
4.4 Effect sizes	150
4.5 Intervention integrity	152
 CHAPTER 5: DISCUSSION	 154
5.1 Introduction to the discussion chapter	154
5.2 Discussion of results and data analysis	154
<i>5.2.1 New Beginnings</i>	154
<i>5.2.2 Getting On and Falling Out</i>	157
<i>5.2.3 Issues of data analysis</i>	159
5.3 Discussion of results in relation to the research questions	160
<i>5.3.1 The effects on children's social and emotional outcomes</i>	160
<i>Pupil self-report</i>	161
<i>Parent ratings</i>	161
<i>Teacher ratings</i>	163
<i>Medium-term follow-up ratings</i>	164
<i>5.3.2 The effects of the intervention on children's self-esteem</i>	164
<i>5.3.3 The fidelity to government guidance on running the intervention</i>	166
5.4 Discussion of methodological implications and limitations	167
<i>5.4.1 Threats to internal validity</i>	167
<i>5.4.2 Threats to external validity</i>	169
<i>5.4.3 Limitations of the study</i>	171
5.5 Discussion of results in relation to existing research findings	173
5.6 Strengths of the study	177
5.7 Professional implications of the research	180
5.8 Future research	182
5.9 Personal learning points	184
5.10 Conclusion	186
 REFERENCES	 187
 APPENDICES	 199

LISTS OF TABLES AND FIGURES

TABLES

<i>Table 2.1: Examples of skills for each process in Salovey and Mayer's (1990) model of EI</i>	24
<i>Table 2.2: The results of the first systematic literature search</i>	48
<i>Table 2.3: The results of the second systematic literature search</i>	49-50
<i>Table 2.4: The results of the third systematic literature search</i>	51
<i>Table 3.1: Threats to internal validity</i>	87
<i>Table 3.2 Outcome measures used in the study</i>	95
<i>Table 3.3: Main researcher activities and when they occurred</i>	101
<i>Table 3.4: The number of children in the intervention and comparison groups</i>	110
<i>Table 3.5: The number of children from each year group participating in the study</i>	111
<i>Table 3.6: The number of boys and girls participating in the study</i>	111
<i>Table 3.7: The number of children from different ethnic groups participating in the study</i>	111
<i>Table 3.8: The number of children speaking English as an additional language participating in the study</i>	111
<i>Table 3.9: The number of children with special educational needs participating in the study</i>	112
<i>Table 3.10: The reasons including the children in the interventions</i>	112
<i>Table 4.1 The sub-scales and informants for each instrument used in the study</i>	122
<i>Table 4.2: The results of the tests of the distribution of the data for the New Beginnings evaluation with non-normal results highlighted in bold</i>	127
<i>Table 4.3: The results of the tests of the distribution of the data for the Getting On and Falling Out evaluation with non-normal results highlighted in bold</i>	128
<i>Table 4.4 The mean, median, standard deviation and range of scores for the New Beginnings evaluation at pre-test</i>	129

<i>Table 4.5 The mean, median, standard deviation and range of scores for the New Beginnings evaluation at post-test</i>	130
<i>Table 4.6 The mean, median, standard deviation and range of scores for the New Beginnings evaluation at follow-up</i>	131
<i>Table 4.7 The mean, median, standard deviation and range of scores for the Getting On and Falling Out evaluation at pre-test</i>	132
<i>Table 4.8 The mean, median, standard deviation and range of scores for the Getting On and Falling Out evaluation at post-test</i>	133
<i>Table 4.9: Median scores in the Getting On and Falling Out evaluation, for the dependent variables that were significantly different at pre-test but not at post-test</i>	137
<i>Table 4.10: The results of the Wilcoxon Signed Ranks Tests for pre- to post-test change (New Beginnings)</i>	142
<i>Table 4.11: The results of the Wilcoxon Signed Ranks Tests for pre- to post-test change (Getting On and Falling Out)</i>	144
<i>Table 4.12: Effect sizes for the significant changes for pre- to post-test scores for the intervention groups</i>	151
<i>Table 4.13: Effect sizes for intervention groups' significant changes from pre-test to post-test for sub-groups, and for pre-test to follow-up gains</i>	152
<i>Table 4.14: Fidelity to the intervention in the six schools</i>	153

FIGURES

<i>Figure 2.1: The waves of intervention model (DfES, 2005)</i>	44
<i>Figure 3.1: The chosen research design</i>	90
<i>Figure 3.2: The design, timescales and interventions</i>	102
<i>Figure 3.3 The procedure and outcomes for recruiting schools to the study</i>	104
<i>Figure 4.1: The design, timescales and interventions</i>	124
<i>Figure 4.2 The differences between the intervention and control groups' scores at pre- and post-test on parent-rated empathy (Getting On and Falling Out)</i>	138
<i>Figure 4.3 The differences between the intervention and control groups' scores at pre- and post-test on teacher-rated empathy (Getting On and Falling Out)</i>	139
<i>Figure 4.4 The differences between the intervention and control groups' scores at pre- and post-test on teacher-rated behaviour difficulties (Getting On and Falling Out)</i>	139
<i>Figure 4.5 The differences between the intervention and control groups' scores at pre- and post-test on teacher-rated pro-social behaviour (Getting On and Falling Out)</i>	140

GLOSSARY OF KEY ABBREVIATIONS USED

DCSF – Department for Children, Schools and Families

DfES – Department for Education and Skills

EI – Emotional intelligence

ELAI – Emotional Literacy Assessment Instrument

EP – Educational psychologist

RCT – Randomised controlled trial

SDQ – Strengths and Difficulties Questionnaire

SEAL – Social and emotional aspects of learning

SEL – Social and emotional learning

SEN – Special educational needs

CHAPTER 1: INTRODUCTION

1.1 The context of and rationale for the research

1.1.1 *The initial training of educational psychologists*

This section highlights influences on the study, including university, local authority and Development and Research (D and R) project requirements.

The present study was completed to meet the requirements of a Doctorate in Applied Educational Psychology at the University of Nottingham. This degree also serves as the professional qualification for new entrants to the educational psychology profession in the UK, having replaced the one year masters course in 2006. In the second and third years of the training course, trainee educational psychologists (EPs) spend most of their time on placement in a local authority and in many cases remain in the same local authority for a two year period. Trainees complete a research project during this time, so it follows that many choose a research area of benefit to their local authority.

Another change has been the D and R project; a collaborative endeavour between some training courses that aims to co-ordinate trainees' research on priorities identified by Principal EPs. The priorities for 2006 to 2009 were tackling bullying, including children with autism, targeted interventions to prevent exclusion and the effects of parent training programmes. In 2007 the areas were broadened to cover developing psychological wellbeing,

promoting social inclusion and enhancing educational outcomes. The overall aim is to strengthen evidence-based practice in educational psychology (see section 1.1.2) through aggregating the results of trainee EPs' research. It means that studies must be described systematically and the Strengths and Difficulties Questionnaire must be used as an outcome measure.

1.1.2 Evidence-based practice and the focus on children's outcomes

Another influence is evidence-based practice (EBP), which is part of a government agenda to increase fairness and consistency in public services by encouraging professionals to base decisions on good quality research evidence rather than personal beliefs (Fox, 2002). This concept of 'quality' research has led to the acceptance of a research hierarchy "based on a logical positivist view of reality," (Fox, 2002, page 43) with systematic reviews and randomised controlled trials at the top of the hierarchy and personal opinion at the bottom. Fox (2002) suggests that research from a different epistemological position has little influence and value in this context.

EBP in educational psychology focuses on consuming and producing research on EPs' interventions. Frederickson (2002) writes that more research is needed on psychological interventions for children and that "much of the onus for broadening the defensible evidence base for practice in educational psychology must rest with the profession," (page 101). She implies that EPs may be sidelined if this does not occur, by suggesting that if there continues to be little evidence for EPs' work then alternative providers may be favoured.

Frederickson (2002) writes that evaluating outcomes is at the core of EBP. This links with another political agenda, focusing on outcomes for children, arising from Every Child Matters (ECM, DfES, 2003). ECM has five outcomes for professionals to work towards: being healthy, staying safe, enjoying and achieving, making a positive contribution and economic wellbeing. This has arguably created a climate where EPs are under pressure to demonstrate how their work improves children's outcomes in these areas.

The current climate of EBP and outcomes for children means that research into psychological interventions for children is needed to demonstrate EPs' value within the children's workforce. It steers the researcher towards a more positivist position, as EPs seek to show that recommended interventions are based on the best available evidence in the research hierarchy.

1.1.3 Meeting children's emotional needs in schools

ECM has also broadened the role of school staff. For example schools must demonstrate how they contribute to meeting the five outcomes when they are inspected, rather than being judged purely on academic measures (Ofsted, 2010). A parallel development has been a focus on emotional wellbeing in schools. For example the Healthy Schools award has a strand on emotional health and wellbeing (Department of Health and DCSF, 2007).

However, the largest scale intervention of recent years to promote children's social and emotional skills in schools has been the Social and Emotional

Aspects of Learning (SEAL) resource. SEAL is a government-designed programme to improve behaviour and attendance in schools. There are Primary and Secondary versions, but both aim to improve children's social and emotional skills by creating a favourable school ethos and using curriculum materials to provide direct teaching opportunities. SEAL is recommended to schools nationwide. The Primary SEAL guidance document gives advice to schools on implementing and evaluating SEAL. However, it could be argued that assessing children's social and emotional skills is an unfamiliar area for many members of school staff, and is something that EPs could assist with.

The introduction of SEAL has led to media scepticism, for example the Daily Mail website (no date) dubs SEAL as 'happiness lessons' that leave less time for traditional subjects. There is a hint of political scepticism at schools broadening their remit, for example the title of the Department for Children, Schools and Families was changed to the Department for Education within hours of a new UK government forming, implying that schools should focus on traditional learning. SEAL has also had a mixed reception in academia, for instance Ecclestone and Hayes (2009) challenge schools' focus on emotions. SEAL would appear to be an area in need of research.

1.1.4 Personal interest

Although the preceding sections describe the influences and constraints on the researcher, it is important to point out that the research topic is also an area of personal interest. Having taught in inner city primary schools pre-

SEAL, the researcher observed the effects of social and emotional difficulties on children's enjoyment of school and attainment. For example, it was often necessary to help children with managing playground conflicts, tolerating frustration, feeling in control of their learning, managing anger and working in groups. However, with a crowded timetable, this was often done as situations arose. The researcher was very interested to find out more about a pro-active and structured approach to help children develop skills to manage the social and emotional aspects of their learning.

1.2 Aim of the study and overview of the chapters

The primary aim of this study was to find out about the impact of SEAL on children's psychological outcomes. The focus was on the small group aspect of SEAL. More information on this is provided in the next chapter.

The following chapters detail the research project. Chapter 2 explains more about the SEAL resource and its theoretical underpinnings, examines the research evidence for SEAL and other similar interventions, and uses this to develop research questions. Chapter 3 details the methodology used to answer the research questions, including the epistemological stance, research design, data collection and data analysis. Chapter 4 presents the results of the study, including statistical analyses. Chapter 5 discusses the results in more detail and in relation to the research questions, methodological issues, existing literature, future research and professional issues.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction to the literature review chapter

This research aims to investigate the effects of the SEAL silver set materials on children's social and emotional competence. The silver set materials are a targeted element of the SEAL programme and details of their contents are given in sections 3.5.3. This chapter provides a research context for the present study by describing the Primary SEAL resource and its development (section 2.2), which reveals that the SEAL initiative draws upon several psychological theories, reviewed in section 2.3. SEAL is also based upon North American practice known as social and emotional learning (SEL). SEL is defined and its evidence-base explored, in section 2.4. Having set the broader context, section 2.5 narrows the focus to return to the silver set materials, and how they fit into the government's waves of intervention model. Then follows a systematic review of research on the effectiveness of the silver set materials and similar interventions in section 2.6. Previous research on SEAL in Central Bedfordshire is discussed in section 2.7, which leads to a discussion of the gaps in the literature and rationale for the present study (section 2.8). The chapter concludes by stating the research questions that were investigated by the study.

2.2 The historical and theoretical underpinnings of Primary SEAL

2.2.1 Overview of the Primary SEAL resource

Being the government's recommended resource in this area, SEAL is arguably the major way of developing children's social and emotional skills in British schools. Following a pilot period (the Behaviour and Attendance pilot, described in Hallam et al, 2006), Primary SEAL was launched in 2005 (DfES, 2005) followed by Secondary SEAL in 2007 (DfES, 2007b). SEAL is not compulsory, but is recommended to schools in a similar way to the National Strategies for literacy and numeracy; schools can provide for children as they see fit, but must show how this meets key government-determined outcomes (DfES, 2006b).

The Primary SEAL resource aims to provide teachers with a structured whole-school curriculum framework for developing children's social, emotional and behavioural skills (DfES, 2005). It is organised into six themes, each designed to last half a term, and each targeting different skills: New Beginnings (focusing mainly on empathy), Getting On and Falling Out (managing feelings), Going for Goals (motivation), Good to be Me (self-awareness and managing feelings), Relationships (self-awareness and managing feelings) and Changes (motivation and social skills). There is also a shorter theme, Say No to Bullying, which is designed to tie into Anti-Bullying Week.

Each theme is introduced through a whole-school assembly and followed up with curriculum materials for different age groups: Foundation stage (red set),

Years 1 and 2 (blue set), Years 3 and 4 (yellow set) and Years 5 and 6 (green set) to enable children of all ages to follow the same theme simultaneously and to provide a spiral curriculum that gradually develops children's skills over time. The resource also includes ideas for developing staff knowledge and confidence (purple set) and involving parents and carers (gold set). The focus of this research is the silver set materials, which are described in more detail in the methodology chapter. They provide a resource for small group work that gives targeted children a chance to have extra input and the opportunity to practise the same SEAL skills introduced in lessons and assemblies.

2.2.2 Primary SEAL: aims and origins

The DfES (2006b) describes the origins of Primary SEAL, arguing that it was developed in response to a substantial evidence-base from the USA on the impact of SEL on school achievement, amongst other areas. The government is said to have consulted research evidence on effective practice in SEL, such as Wells et al's (2003) systematic review on mental health promotion in schools, Weare and Gray's (2003) research report on what works in developing children's social and emotional competence and Elias et al's (1997) research synthesis from the USA (which are examined in section 2.4.2). From this evidence, key features of effective SEL provision were identified to guide the development of SEAL. One might ask 'why not use the existing evidence-based programmes available rather than creating such a large resource from scratch?' The DfES (2006b) explain that "the aim was to deliver similar content, but in a UK context and with new elements that would achieve

additional outcomes identified by the department for a UK programme,” (page 1). The additional outcomes were based upon government priorities such as school improvement, raising standards, children’s mental health and tackling social exclusion, as it was felt that American programmes focused more on social skills and reducing aggression.

2.2.3 Primary SEAL: theoretical underpinnings

The DfES (2006b) has stated the authors and psychological theories that influenced the content of SEAL. These are: affective competencies (also known as emotional intelligence or emotional literacy, Salovey and Mayer, 1990, and Goleman, 1995), empathy (Feshbach, 1975), social problem solving (Spivack and Shure, 1974), anger management (Novaco, 1976) and cognitive-behavioural theories (Bandura, 1986, and Kendall, 2000). These are examined in section 2.3. It is claimed that SEAL differs from other SEL programmes as it targets all of these areas, rather than only focusing on some of them, as other common programmes do (DfES, 2006b). Another unique feature, according to the DfES (2006b), is that SEAL aims to develop children as learners, by including lessons on motivation, awareness of oneself as a learner, appreciation of multiple intelligences and the skills needed for group work. However, no references are given for these areas, so it is difficult to judge how successfully the resource incorporates the key theoretical messages.

The government argues that SEAL retains features from evidence-based American programmes whilst having intentional differences as well (DfES,

2006b). The differences are stated as: focusing on managing unpleasant feelings other than anger, having an explicit focus on skills needed as a learner such as persistence, having an explicit focus on diversity and anti-bullying, being a less prescriptive and more enquiry-led approach, having a cross-curricular approach, using ICT as a learning medium and having a shared whole school focus on a theme (DfES, 2006b). Sound reasons are given for these differences, including maximising the take-up by schools and fitting with schools' existing priorities. However, the government does not provide research evidence to justify these changes and admits that "the initial evaluation by the Institute of Education is promising, although it needs to be followed up with in-depth and ongoing micro-analysis," (DfES, 2006b, page 4). The reassurance that "the fact that SEAL shares its theoretical basis with evidence-based programmes used overseas, and the robust match between its core features and those established by research as fundamental to success, give comfort to users that impact will be at least as good as that of longer-established programmes," (DfES, 2006b, page 4) is insufficient, in this researcher's opinion, to justify such significant deviation from the research on what works in SEL. A central argument of this thesis is that it is not enough to be comforted by research into the programmes that *inspired* SEAL, and that the SEAL programme *itself* should have been well-evaluated before being rolled out nationwide.

2.3 Psychological theories that influenced the content of SEAL

This section explores the research introduced in section 2.2.3.

2.3.1 Emotional intelligence (EI)

Origins and definitions

In 1990 an article called 'Emotional Intelligence' was published by Salovey and Mayer. They wrote that, despite earlier thinking that emotions are a hindrance to intelligence, emotions can actually be adaptive by directing and prioritising cognitive processes. They conceive of EI as a subset of social or personal intelligence and define it as "the ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this information to guide one's thinking and actions," (page 189). EI focuses on recognising and using emotions, rather than self-concept or appraising others.

Salovey and Mayer (1990) discuss the processes and skills involved in EI. These are summarised in table 2.1. They suggest that these skills are subject to individual differences, may be educable, contribute to positive mental health and that a deficit in skills may lead to problems in adjustment such as becoming ostracised, unfulfilled or depressed.

Process	Examples of skills
Appraising and expressing emotion in the self	<ul style="list-style-type: none"> ▪ using language to name and describe emotions ▪ giving non-verbal signals such as facial expressions
Appraising and expressing emotion in others	<ul style="list-style-type: none"> ▪ interpreting non-verbal expressions ▪ empathy
Regulating emotion in the self	<ul style="list-style-type: none"> ▪ using behaviour such as choosing who to spend time with or engaging in altruistic acts to maintain or change moods
Regulating emotion in others	<ul style="list-style-type: none"> ▪ creating a favourable impression of oneself with others
Using emotions	<p>Emotions can aid problem solving by helping us to:</p> <ul style="list-style-type: none"> ▪ consider more alternatives ▪ be more organised ▪ focus ▪ prioritise ▪ motivate ourselves

Table 2.1: Examples of skills for each process in Salovey and Mayer's (1990) model of EI

Popularisation

Goleman (1995) adapted Mayer and Salovey's model and brought this academic theory to the mainstream. Goleman's model features five social and emotional competencies: self-awareness, self-regulation, motivation, empathy and social skills. SEAL is based upon an identical model (DfES, 2005).

Goleman's work is somewhat notorious for its claim that EI is more influential in predicting life success than intelligence quotient (IQ). Goleman (1995) writes that EI can enhance intellectual potential; a person with high EI can outperform someone else with a higher IQ but lower EI. Goleman (1995) cites longitudinal studies where IQ did not predict life success and concludes that EI

could be more influential than IQ and, unlike IQ, is not fixed. Goleman (1995) asserts that EI prepares people to cope with adversity. He appears to implicate lack of EI in serious social problems such as gun crime and despairs that schools prioritise teaching academic, over emotional, skills.

Controversy and fragmentation

Goleman's (1995) landmark book has attracted scorn from within academic psychology (for example Matthews et al, 2004b, Mayer et al, 2000b and Ashkanasy and Daus, 2005). Lack of definition appears to be the major criticism. Rather than restricting EI to emotional abilities, Goleman's model also includes character traits, expanding the concept to include aspects of personality. The main problem is that this implies that these features are connected or form a package, when they might be unrelated (Craig, 2007).

Another problem with the loose definition is the possibility that EI is a repackaged construct. For example, Qualter, Gardner et al (2007) highlight similarities between EI and social cognition, social information processing and emotion regulation. They argue that correlations between EI and personality are inconsistent. They also examine the link between EI and intelligence and conclude "we argue that [trait and ability EI] should be seen as umbrella terms, encompassing many previously investigated and empirically supported psychological constructs," (page 12). In other words, Goleman's (1995) ideas are a rebranding of existing personality and intelligence research.

Goleman's (1995) text appears to have led to two models of EI. Qualter, Gardner et al (2007) suggest ability models view EI as the ability to process information about emotions, and are thus linked with intelligence, whereas trait models involve self-perceptions and dispositions linked with emotion so have more in common with personality. This implies that Goleman's EI should not be classed as an intelligence. The varying definitions have led to diverse approaches to EI's measurement. Ability EI is usually measured through performance, whereas trait EI is often measured by self-report.

The application of EI to education has been criticised too, for example Waterhouse (2006) writes that this should not happen until there is more evidence. Mayer and Cobb (2000) argue that the education world's take-up of Goleman's view of EI is due to policy-makers being informed by journalism, rather than scientific evidence. Craig (2007) writes that "Goleman's work is also at the core of both Primary and Secondary SEAL... [it] cannot credibly be used as the intellectual foundation, and justification of large-scale work of this type in school," (page 24). However, although Goleman's five domains are used, SEAL does not claim to be purely an EI intervention.

Not all criticism of EI is of Goleman's work however. Craig (2007) argues that the "intellectually respectable end" (page 8) of EI has been criticised and, since it is a relatively new area, is in need of further research. Salovey and Mayer (1990) defend their construct against the accusation that it concerns unrelated component skills. They argue that, although the parts may or may

not be inter-correlated, they all meet their criteria of involving emotional processing and being necessary for a minimum level of competence.

The future?

Despite the criticism of Goleman's work, even his opponents acknowledge that some aspects of the theory are desirable. For example, Craig (2007) writes that even though Goleman overstates the importance of EI, it "does not mean that it does not matter at all," (page 15) and goes on to describe the importance of optimism, flow and the positive psychology movement.

Matthews et al (2004a) discuss seven myths about EI, which summarise the main criticisms. They do not reject EI but propose further research to support or disconfirm the myths. They rate the future prospects of EI, with poor prospects for self-report measures and linking EI to emotional functioning but fair or good chances for the others. This suggests a field in its early stages and in need of further empirical research, rather than a field to be dismissed.

Although influenced by EI, practitioners evade controversy over definitions by avoiding the term altogether. Applied psychologists often use the term emotional literacy. The next section discusses this term and its relation to EI.

2.3.2 Emotional literacy

Qualter, Whiteley et al (2007) use EI and emotional literacy synonymously and state that the latter term is preferred in the UK. Sharp (2001) suggests that

Claude Steiner coined the term emotional literacy. He defines it as “the ability to recognise, understand, handle, and appropriately express emotions... to help yourself and others succeed,” (Sharp, 2001, page 1). Tew (2007) differentiates EI and emotional literacy, arguing that EI is an ability whereas the latter involves skills, attributes and competences and their development.

Weare (2004) defines emotional literacy at the individual and organisational levels. For individuals, her definition is very similar to the one given by Sharp (2001). At the organisational level, it is “the extent to which the organization takes into account the role of emotion in dealing with the people who are its members, and in planning, making, and implementing decisions, and takes positive steps to promote the emotional and social well-being of its members,” (page 3). SEAL incorporates this by taking a whole school approach and including materials on staff wellbeing. Weare (2004) lists key competences involved in emotional literacy, which match very closely with Goleman’s (1995) model of EI.

Weare and Gray (2003) cite some reasons for using the term ‘literacy’ rather than ‘intelligence’ in relation to emotional and social skills. For example, EI might not meet the criteria for an intelligence, intelligence has connotations of measurement and fixedness and, when used loosely, EI has a similar meaning to emotional literacy or social competence, so no specialist meaning is added. However, they also present disadvantages, for example ‘emotional literacy’ can be taken as a within-child concept and minimise social aspects.

They argue that the terms 'emotional and social competence' and 'emotional and social wellbeing' are preferable, with the former relating to learned knowledge and skills and the latter referring to environmental determinants.

2.3.3 Empathy

Feshbach (1975) defines empathy as "a match between the affective response of a perceiver and that of a stimulus person," (page 26). She differentiates it from sympathy (where a person understands another's emotional state) and projection (where the observer's characteristics are attributed to another person) by arguing that "the subject assumes the emotional attributes of the stimulus person," (page 25). Feshbach (1975) suggests that empathy has cognitive and affective components, and comprises three key abilities: discriminating the perspective of another person, discriminating the role of another person and capacity for emotional responsiveness.

Feshbach (1975) describes several studies conducted in the late 1960s and early 1970s using the Affective Situation Test, that were designed to find out more about empathy in children. One finding was that children were more empathic when the 'stimulus child' was perceived to be similar to them, for example in terms of gender and race. It was also found that children become more empathic between the ages of five and eight, and that girls are more empathic than boys between the ages of four and seven (however, the findings on sex differences were less consistent than other results).

The relationship between empathy and aggression has been investigated. Feshbach (1975) proposes that empathy inhibits aggression, as the aggressor would experience the victim's emotional pain. The hypothesised negative correlation between the two was supported by several studies of older boys and male college students. However, for younger boys, those classed as being high in empathy were actually *more* aggressive. Feshbach (1975) argues that, in common with similar studies, aggression in preschoolers can be a reflection of maturity rather than hostility. There was no relationship between empathy and aggression in girls of any age, which Feshbach (1975) attributes to the limited range of girls' aggression.

The applications of empathy are discussed in Feshbach's (1975) article. For example, the construct could be used to inform training programmes to promote pro-social behaviour, by using role play techniques and by maximising perceived similarity between people. The influence of Feshbach's (1975) research can be seen in the SEAL programme, as empathy is one of the five skills it emphasises. It is interesting that the government chose to focus on the positive aspect (promoting empathy), rather than the negative (reducing aggression), unlike US approaches to SEL. One criticism might be the lack of empirical support for the relationship between aggression and empathy in girls and younger boys, implying that targeting empathy may not reduce aggression for these groups.

2.3.4 Social problem solving

Spivack and Shure (1974) developed a training programme to teach cognitive interpersonal problem solving skills to targeted four year olds. It consists of scripted daily lessons, in the form of games, designed to be presented to small groups by teachers. There are forty six lessons, each lasting between five and twenty minutes, so the programme typically takes at least nine weeks to deliver. The authors state that it is unnecessary to follow the script exactly and leaders can complete more or less than suggested in one day, depending on the restlessness of the group. The daily sessions progress from word concepts to pre-problem solving skills to interpersonal problem solving skills.

The authors describe three evaluation studies (two preliminary and one comprehensive) of the programme, conducted whilst it ran in Head Start centres in Philadelphia, USA, in the early 1970s. The studies are progressively large in scale, with around 20 children per condition in the first, and around 100 per group in the third. The training group in the first study was led by research assistants, and the other studies used the children's teachers.

In the first evaluation, the training group's improvements in interpersonal problem-solving were significantly better than those in the attention and no treatment control groups at post-test. Class teachers' behaviour ratings indicated that the children who were classified as behaviourally 'aberrant' at pre-test were more likely to improve their ability to delay gratification if they had the training programme, and this difference was statistically significant.

Similarly, in the second evaluation, children in the training group improved their interpersonal problem solving skills compared with matched controls, although exact figures are not given. For behaviour ratings, 52% of the intervention group and 60% of the control group were classified as adjusted at pre-test, whereas at post-test, 85% of the intervention group and 57% of the control group were classified this way. New findings were that teachers were trained to successfully use the programme and that children with the lowest problem solving scores at pre-test made the biggest improvements as a result of receiving the training.

In the third study, the training groups improved their ability to: think of multiple alternative solutions to problems, see different consequences to hypothetical acts and see causal connections in interpersonal events, with the largest improvements for children categorised as 'aberrant' at pre-test. Children in the training groups reduced the number and priority of forceful solutions to interpersonal problems, compared with the control group which increased on both. For behavioural adjustment at post-test, 50% of the training group who were originally categorised as impulsive and 75% classed as inhibited had moved into the adjusted category (compared with 21% and 35% of the control group respectively). 90% of the training group and 86% of the control group that were rated as adjusted at pre-test were also rated as adjusted at post-test. At six month follow-up (when the children had moved from preschool to kindergarten) the authors found that, of the children rated

as adjusted at post-test, 86% of the training group remained so, compared with 66% of the control group. Of the children in the training group rated as aberrant at pre-test and adjusted at post-test, 30 out of 36 remained adjusted at follow-up. There were insufficient numbers of control group children in this category to permit statistical comparison. However, for children rated as adjusted at pre- and post-test, the proportion remaining adjusted at follow-up was significantly higher in the training group than the control group.

In conclusion, Spivack and Shure (1974) argue that the programme leads to improvements in interpersonal problem solving skills and behavioural adjustment across the IQ range and for children initially rated as adjusted, inhibited and impulsive. The gains appear to endure across time, raters and settings. The authors contend that the gains in behaviour occur as a function of improved interpersonal problem solving skills. Whilst the evaluation lacks the detail that was presumably included in preceding journal articles, it provides a useful basis for developing the SEAL materials. For example, it shows that young children in educational settings can be taught social problem solving skills, highlights the link between these skills and positive behaviour and underscores the importance of evidence-based interventions.

A recent review of social and emotional curricula for young children (Joseph and Strain, 2003) commends a later version of this programme for assessing treatment generalisation, treatment maintenance, replication across studies, having evidence for racially diverse groups and replication across settings. It

rates the evidence for the programme as medium. However, it argues that using hypothetical problems and behaviour ratings (rather than direct measures), are limitations of the research. Also, the programme did not meet the reviewers' criteria for assessing treatment fidelity, social validity of outcomes, acceptability of intervention and replication across clinical groups.

2.3.5 Anger management

Novaco (1975) describes the evaluation of a programme for the treatment of anger problems. He notes that anger can have positive functions, therefore the treatment aims to improve individuals' management of their anger rather than eliminate it entirely. The programme uses techniques previously used in the treatment of anxiety: self-instruction and relaxation training. Self-instruction is a cognitive-behavioural technique (see section 2.3.6) that aims to change a person's behaviour by changing their self-statements.

Novaco (1975) conducted an experiment with four conditions to evaluate the treatment programme: self-instruction combined with relaxation training, self-instruction alone, relaxation training alone and an attention control condition. 34 participants (university staff and students and local residents, with a median age of 22) were recruited by responding to an advertisement. All had chronic anger problems, as assessed by pre-tests and a diary kept during the intervention. Novaco states that participants were stratified for gender and randomly assigned to conditions. Participants in the attention control condition were told their experience was part of an extended

intervention with a pre-treatment phase, and were asked to keep a diary of their anger, use rating scales and visit the clinic to discuss the results. Participants were assessed before and after the intervention using an anger inventory (90 hypothetical incidents where participants had to rate how angry they would be on a five point scale) and laboratory provocations. In the laboratory, the measures were self-report of anger, blood pressure, galvanic skin response and a coping strategy measure. The provocations were imaginary or role play at pre-test and imaginary, role play and direct at post-test (with the deception revealed after the direct provocation).

Results were analysed from pre- to post-test. On the anger inventory, all groups decreased their scores but the combined intervention and self-instruction were significantly better than the control condition. There was no significant difference between the relaxation training and the control group. In response to laboratory provocations, there were four measures. On the self-report measure, participants in the combined condition reduced their anger but this was not significantly better than the single intervention conditions. All intervention groups were significantly different from the attention control group, but only on certain scenarios. On the blood pressure measures, there were no significant differences between groups for diastolic blood pressure and, for systolic blood pressure, the combined condition showed significant improvements over the control group on three out of four provocations. The results for galvanic skin responses were less impressive, with no significant differences between the combined or relaxation conditions

and the control group and only one out of four significant results for self-instruction compared with the control group. There were similar equivocal results for coping strategy measures. The final measure was conducted only at post-test, due to the deception involved, and was a direct provocation in the laboratory. The results are not straightforward, since almost half of participants guessed what was happening. There were some positive results in the combined condition for blood pressure and coping strategies.

Novaco (1975) concludes that “without doubt, the combined treatment condition resulted in a very significant improvement in subjects’ ability to regulate and manage anger in comparison to the attention control condition,” (page 43). There was also some support for the self-instruction treatment, and support for the relaxation training but only on certain provocations, suggesting limited generalisability. The results suggest that cognitive and relaxation techniques can help individuals to manage their feelings of anger. However, a criticism of applying this research to SEAL is that the treatment was developed for adults with chronic anger problems, and was not designed for whole school populations of children aged between 4 and 11 years.

2.3.6 Cognitive-behavioural theories

Another theoretical basis for the SEAL programme cited by the DfES (2006b) is the cognitive-behavioural approach. Texts by Bandura (1986) and Kendall (2000) are referred to, and these are briefly summarised below.

Bandura (1986) describes social cognitive theory (SCT), a development of social learning theory. The central idea is 'reciprocal determinism,' whereby environmental events (social), personal factors (cognitive) and behaviour influence each other. Bandura contrasts this with other theories where behaviour is said to be driven by internal forces (psychodynamic and trait theories) or controlled by external stimuli (radical behaviourism).

SCT posits that humans have several basic capabilities. These are: symbolising (using thought and imagination rather than always needing direct experience), forethought (imagining the future, planning, goal setting and anticipating consequences guiding present behaviour), vicarious learning (observational, rather than direct, learning), self-regulation (evaluating behaviour against personal standards to influence future behaviour) and self-reflection (metacognition, including self-efficacy, and evaluation of experiences).

Bandura's work is more theoretical than applied, as it focuses on the mechanisms underpinning thought and behaviour. However it has real world applications, for example in guiding cognitive-behavioural therapy. In the case of SEAL, the influence of SCT can be seen in the domains of emotional literacy. SCT contributes to our understanding of self-awareness, self-management and motivation for example. SCT also influences the delivery of the SEAL curriculum materials, for example by highlighting the importance of teachers modelling strategies to pupils, so that pupils learn vicariously.

The interactionist perspective on behaviour has been influential in British applied educational psychology (for example, see Frederickson and Cline, 2002). In a critique of SCT, Cobb (1992) argues that its strength is that it combines the roles of conditioning, self-reflection and wider socio-political circumstances in guiding human behaviour. She praises SCT for its precisely defined constructs, testability and applications. However, SCT has also been subject to some criticism. For example, Cobb (1992) argues that SCT provides neither normative information nor a lifespan developmental model. Martin et al (2002) argue that SCT lacks crucial detail on how external standards are internalised to become personal standards.

Kendall's (2000) chapter explains the use of cognitive-behavioural therapy (CBT) with children and adolescents. Kendall defines CBT as an amalgam of behavioural procedures and the client's cognitive activities and emotional experiences, which aims to produce changes in thinking, feeling and behaviour. As such, it considers clients' internal and external worlds. Kendall discusses the various roles of a CBT therapist: to collaborate (therapist as consultant), to decode social information (therapist as diagnostician) and to teach (therapist as educator). A principal tenet of CBT, according to Kendall, is that it can challenge and change a person's cognitive structures, processes and products and therefore make their understanding of the world more adaptive. Such cognitive apparatus includes attributions, schemata, beliefs and expectations. Kendall argues that the goal of CBT is to equip individuals with problem solving skills that allow them to cope with life's challenges.

Kendall discusses the particular issues of using CBT with children and young people. He highlights the importance of social context, emphasising the role of peer and family relationships in children's adjustment since children are not yet capable of independence. He also identifies the need to differentiate between cognitive deficiency (lack of thought processes, for example in children with attention deficit disorders) and cognitive distortion (dysfunctional thought processes, for example in children with eating disorders) in order to appropriately direct interventions; in other words, whether to teach new strategies or change existing processes. A final issue is that the therapeutic process should be enjoyable and motivating for children, since they are unlikely to have sought it for themselves.

The fact that Kendall's (2000) work explicitly concerns therapy for children is positive, as much of the CBT literature relates to adults. However, it could also be considered a controversial source of inspiration for SEAL. For example, popular books (such as 'The Dangerous Rise of Therapeutic Education' by Ecclestone and Hayes, 2008) have criticised the government for allowing therapeutic ideas to permeate schools to the detriment of children's development and education. Ecclestone (2007) argues that normalising therapeutic interventions for a mass audience should be resisted, since it leads to "diminished images of human potential and resilience" (page 467).

Another criticism of using this text to create SEAL is that Kendall clearly argues that guiding theories are “necessary but not sufficient” (page 7) and that empirical support for interventions is crucial. Yet the government appears to have used Kendall’s book as evidence for CBT, rather than conducting an evaluation of the SEAL programme before its nationwide rollout to schools.

2.3.7 Summary

The sources of inspiration for the SEAL programme have been examined. These have included experimental psychology research from the 1970s into treatment programmes with good results but varying applicability to SEAL, well established psychological theories and more recent and controversial theories and therapies. Despite these theories and research having some good evidence, the argument remains that this is evidence for the inspiration for SEAL, rather than for the programme itself.

2.4 Educational practice that influenced the development of SEAL

2.4.1 Social and emotional learning: definition and origins

Social and emotional learning (SEL) is a growing area of practice in the USA that has hugely influenced the development of the Primary SEAL resource (DfES, 2006b). The term SEL is used in a similar way to the term emotional literacy in this country. SEL refers to the education of knowledge, skills and attitudes that enable children to recognise and manage emotions, care about others, make good decisions, behave ethically and responsibly, develop positive relationships and avoid negative behaviours (Zins et al, 2003).

SEL appears acceptable to critics of applied EI. For example, Mayer and Cobb (2000) wrote “although we do not think it makes sense to talk about the ready acquisition of emotional intelligence, a slight change in language – to socioemotional learning – is entirely acceptable to us,” (page 177). Their message is that such curricula should be evidence-based and tied to science not journalism. The next section considers the evidence-base for SEL.

2.4.2 Reviews of SEL research that influenced the development of SEAL

The DfES (2006b) is known to have consulted three summaries of research on SEL when developing Primary SEAL, and these are now appraised in turn.

Elias et al (1997), on behalf of the Collaborative for the Advancement of Social and Emotional Learning (CASEL) produced guidance on promoting SEL. The text is practical rather than a review of empirical studies and therefore does not provide a great deal of information on the research methods used to compile it. The book offers 39 guidelines for social and emotional education. These “have a strong scientific basis and are based on many research investigations and relevant theory” (Elias et al, 1997, page 13) as well as the expertise of programme developers, researchers, trainers and practitioners. Each guideline is explained in terms of rationale, implications and applications and supported by examples from the field. Some guidelines are clearly present in the SEAL programme, for example staff training materials and a long-term, developmental approach. Evaluation, the focus of this research

project, is mentioned in several guidelines. Details are given of empirically supported SEL programmes and schools that are delivering them and willing to be contacted. All are in the USA.

Wells et al (2003) conducted a systematic review of universal (rather than targeted) approaches to mental health. Inclusion criteria for the studies were to: be at least partly school-based, aim to promote some aspect of their broad definition of mental health or prevent mental illness, use at least one affective or behavioural measure of mental health, report with adequate detail the content, delivery and outcome measures (including validity and reliability), have a control group that was similar at the start of the study and have a minimum of two groups or 40 individuals with less than 30% attrition. The studies were all written in English. They identified 17 studies that met the criteria, but found that the populations, interventions and outcomes were too diverse to synthesise quantitatively. Four studies found over 70% of positive outcomes on the measures used, five had between 30 and 70% positive outcomes and one study had less than 30% positive outcomes. The remainder had effective results for subgroup analyses only (for example 'white males') rather than universally. The following trends were identified: most studies showed some positive results, the most successful interventions were mental health promoting (rather than mental illness preventing) and lasted over a year, there was support for whole-school approaches (although this was limited to two studies) and most studies were carried out in the USA (all but two).

The government commissioned Weare and Gray (2003) to write a paper on how children's social and emotional competence could be developed at a national and local authority level. They carried out a literature review, interviews and case studies and recommended: developing common terminology (they favour emotional and social competence and wellbeing), finding an appropriate strategic location for the work, developing the evidence-base, promoting the benefits of this work, prioritising it, taking a holistic approach, ensuring coherence (of multi-professional work and involving families and communities), starting early and taking a developmental approach, creating appropriate environments, including explicit teaching and learning programmes and promoting teachers' competence and wellbeing. They recommended that the government provide curriculum guidance to schools, using principles of effective programmes identified in the literature. This paper was clearly very influential in the development of the SEAL materials.

In summary, the three reviews consulted by the government were generally positive about SEL. However, the evidence was based more upon case studies for Elias et al (1997) and Weare and Gray (2003), rather than rigorous evaluations. Also, the evidence mainly relates to SEL in American schools.

2.5 The 'waves of intervention' model of meeting additional needs

Having discussed SEAL on a whole-school or -class basis, this section focuses on the silver set, in the context of the waves of intervention model.

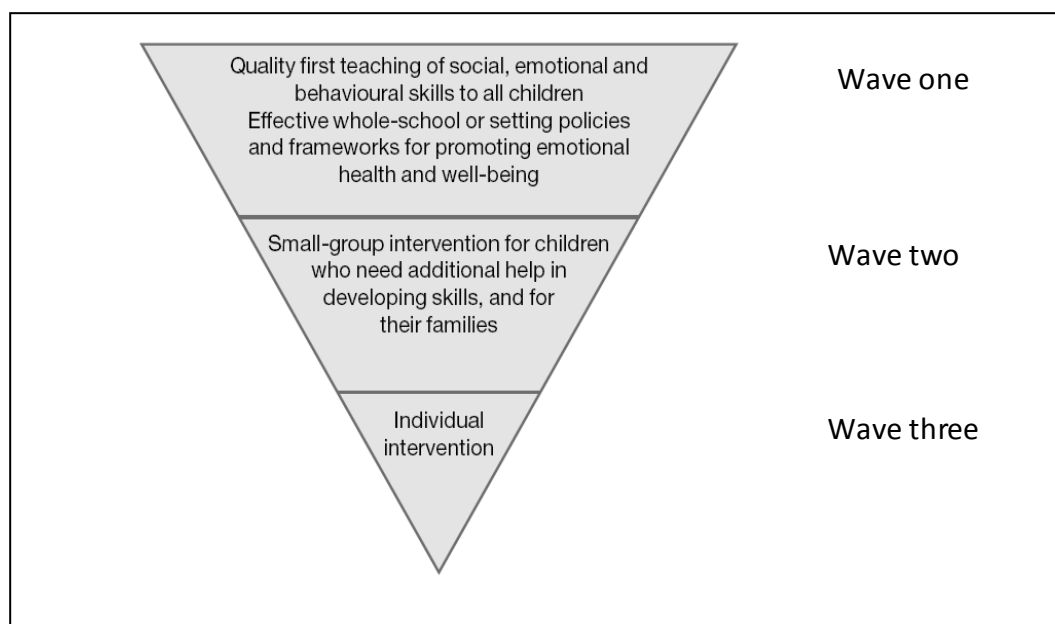


Figure 2.1: The waves of intervention model (DfES, 2005)

Figure 2.1 is a diagram of the model. The idea comes from the government's National Strategies documentation, particularly in relation to literacy and mathematics but also as applied to behaviour (DfES, 2006c). Wave one encompasses quality first teaching and whole-school policies experienced by all pupils, as shown by the large end of the triangle. A poor response to one wave of intervention would lead to a child being provided with the next level of intervention in addition, as shown by the narrower parts of the triangle. The DfES (2002) aligns the waves of intervention with the SEN Code of Practice. The waves of intervention model also relates to a health model of universal or targeted approaches. Humphrey et al (2008) compare wave one with universal intervention and wave two with targeted intervention.

Humphrey et al (2008) discuss how SEAL relates to the waves of intervention model. Wave one SEAL is the development of a whole school ethos and climate that supports social and emotional skills. It also includes quality first teaching of social and emotional skills using the SEAL curriculum materials. Wave two is a small group intervention for targeted children who need further support with their social and emotional development, and the silver set materials are available to schools for this. The aim is to facilitate personal development, explore issues in more depth, practise skills in a safe setting and promote reflection. Wave three is one to one intervention for children with higher level needs, including children with mental health difficulties. The government has commissioned evaluations of SEAL at wave one (Hallam et al, 2006), wave two (Humphrey et al, 2008) and wave three (Wolpert et al, ongoing, cited in Humphrey et al, 2008).

This staged approach to meeting children's social and emotional needs at school is an interesting, and perhaps distinguishing, feature of the silver set SEAL intervention. It assumes that the intervention occurs in a context where actions have been taken to address systemic factors such as school ethos, policies and staff understanding and where staff know that children have already had prior experience of the direct teaching of social and emotional skills. Government guidance refers to two modes of transmission of social and emotional skills: being 'caught or taught' (Department of Health/Department for Education and Skills, 2004). This highlights the idea

that SEL relies on environmental change as well as teaching. So, if the waves of intervention process is adhered to by addressing contextual issues first, it could distinguish wave two SEAL from other similar but standalone interventions. The danger of the latter interventions is the implied position (which is demonstrated through a focus on children and their perceived lack of skill, rather than a systemic focus) that the cause of social and emotional difficulties lies within the child and not the school or educational system.

Of course there is a possibility that, with the silver set materials freely available online, variable levels of local authority involvement in SEAL over time and the introduction of new educational initiatives which reduce the focus on wave one SEAL, the silver set materials are used as a standalone intervention. This could lead to the type of problems described in the previous paragraph and to schools ignoring the advocated focus on the organisation (Weare, 2004), creating appropriate learning environments and promoting teacher wellbeing (Weare and Gray, 2003) and skills being passed on by being caught *and* taught (Department of Health/Department for Education and Skills, 2004).

2.6 Systematic reviews of research into SEL interventions

This section aims to locate the present evaluation study within the existing research context in as unbiased way as possible. Three systematic searches were conducted to achieve this aim. Systematic literature reviews reduce bias by being open about the terminology used in searches and by having clearly stated and objective criteria to select the studies for inclusion in the review. This ensures that the author considers all studies that meet a defined set of criteria, rather than focusing on studies to suit a line of argument. The reason for doing three searches was the small number of studies identified by each.

2.6.1 Details of the systematic search strategies

The databases searched were those available through the University of Nottingham library: ASSIA, ERIC, Intute Education Gateway, PsychInfo, Teacher's Reference, UNLOC, Web of Science and Zetoc. The inclusion criteria for studies in the first search were:

- Studies examining the effectiveness of the DCSF SEAL programme
- Studies from the UK

The results are shown in Table 2.2.

Search term	Number of hits	Action taken	Number of articles meeting inclusion criteria
“silver set”	2	Abstracts/descriptions consulted and inclusion criteria applied	0
SEAL	1441	Search refined	Not applicable
“social and emotional aspects of learning”	33	Abstracts/descriptions consulted and inclusion criteria applied	0
Total studies meeting inclusion criteria:			0 ¹

Table 2.2: The results of the first systematic literature search

Since the searches did not generate any studies, and the SEAL programme is the key area for this study, the search was widened to include articles available online, rather than in journals or books. A search using the Google Scholar website and the search term “social and emotional aspects of learning” resulted in over 300 hits and appeared to include many unrelated articles, so was refined to “social and emotional aspects of learning” +SEAL +evaluation. This search led to around 150 hits. All were examined according to the inclusion criteria, and five met the criteria. These five studies are discussed in section 2.6.2.

¹ This search was repeated immediately prior to the submission of this thesis and uncovered journal articles by Hallam (2009), Humphrey et al (2009) and Lendrum et al (2009). However, these were not included in the review, since they were based on the same data and yet had less detail than the research reports found in the web search and discussed in section 2.6.2.

The inclusion criteria for studies in the second search were:

- Evaluations, or systematic reviews of evaluations, of interventions
- Studies that concerned targeted, rather than universal, interventions
- Studies that aimed to affect social and emotional competence, as evidenced by assessing social or emotional skills, pro-social or problematic behaviour, mental health or mental health difficulties (or perceptions of these), using quantitative or qualitative methods
- Studies conducted with children primarily in an educational setting
- Studies from the UK
- Studies written in English

The search terms and results of the searches are shown in Table 2.3 below.

Where a search generated over 100 results, the search term was narrowed.

Search term	Number of hits	Action taken	Number of articles meeting inclusion criteria	References for studies
"social and emotional learning"	520	Search refined	Not applicable	
"social and emotional learning" +schools	243	Search refined	Not applicable	
"social and emotional learning" +schools	68	Abstracts/descriptions consulted and inclusion criteria applied	0	

+evaluation				
“social and emotional learning” +schools +group	67	Abstracts/descriptions consulted and inclusion criteria applied	0	
“emotional literacy”	281	Search refined	Not applicable	
“emotional literacy” +group	20	Abstracts/descriptions consulted and inclusion criteria applied	1	Sharp and Herrick (2000)
“emotional intelligence” +school	664	Search refined	Not applicable	
“emotional intelligence” +school +evaluation	42	Abstracts/descriptions consulted and inclusion criteria applied	1	Coppock (2007)
“social skills” +groups	3408	Search refined	Not applicable	
“social skills” +groups +school	1121	Search refined	Not applicable	
“social skills” +groups +school +evaluation	203	Search refined	Not applicable	
“social skills” +groups +school +evaluation +intervention	70	Abstracts/descriptions consulted and inclusion criteria applied	2	Maddern et al (2004) Parton and Manby (2009)
“mental health” +school	2253	Search refined	Not applicable	
“mental health” +school +“early intervention”	208	Search refined	Not applicable	
“mental health” +school +“early intervention” +evaluation	15	Abstracts/descriptions consulted and inclusion criteria applied	0	
Total studies meeting inclusion criteria:			4	

Table 2.3: The results of the second systematic literature search

Four studies were included and are discussed in section 2.6.3. Due to the small number of studies that met the criteria, the searches were repeated without the inclusion criterion “studies that concerned targeted, rather than universal, interventions.” This opened up the review to include studies of universal interventions. The results are shown in Table 2.4.

Search term	Number of hits	Action taken	Number of articles meeting inclusion criteria	References for studies
“social and emotional learning” +schools +evaluation	68	Abstracts/ descriptions consulted and inclusion criteria applied	0	
“social and emotional learning” +schools +group	67	Abstracts/ descriptions consulted and inclusion criteria applied	0	
“emotional literacy” +group	20	Abstracts/ descriptions consulted and inclusion criteria applied	1	Matthews (2004)
“emotional intelligence” +school +evaluation	42	Abstracts/ descriptions consulted and inclusion criteria applied	2	Curtis and Norgate (2007) Kelly et al (2004)
“social skills” +groups +school +evaluation +intervention	70	Abstracts/ descriptions consulted and inclusion criteria applied	0	
“mental health” +school +“early intervention” +evaluation	15	Abstracts/ descriptions consulted and inclusion criteria applied	0	
Total studies meeting inclusion criteria:			3	

Table 2.4: The results of the third systematic literature search

Three studies were included in the review and are discussed in section 2.6.4.

2.6.2 Evaluations of SEAL

The systematic search of evaluations of the SEAL programme led to five studies. One was a government-commissioned study to examine the wave two Primary SEAL silver set materials (Humphrey et al, 2008). This is the most relevant to the present research project and is therefore described in the most detail. Another was a conference presentation of the results of a study evaluating the Family SEAL (gold set) materials (Downey and Williams, 2009), which is also positioned at the wave two level. Another government-commissioned study examined the Primary Behaviour and Attendance pilot, which included the wave one SEAL curriculum materials and wave two small group work using other materials than the silver set materials (Hallam et al, 2006). One study was a case study of a school adopting Primary SEAL, with some information about its evaluation (DfES, 2007a). The final study was an evaluation of the social, emotional and behavioural skills (SEBS) pilot, which was a precursor to Secondary SEAL (Smith et al, 2007).

Evaluations of wave two Primary SEAL

Two studies of wave two Primary SEAL were located, one on the silver set materials for targeted children and one on the gold set materials for parents.

Humphrey et al's (2008) evaluation of wave two primary SEAL aimed to investigate the impact of small group work on social and emotional skills, the impact on other areas (such as behaviour, attendance, learning and on schools and families), the longevity of the effects, the implementation of SEAL group work and how it complements wave one SEAL. They used interviews, a quantitative impact evaluation and case studies to meet these aims. Focusing on the impact evaluation, which bears most relevance to the present study, a pre-test-post-test quasi-experimental design was used to evaluate the New Beginnings and Going for Goals interventions and a single-group phase-change design was used to examine Getting On and Falling Out and Good to be Me. The sample size was relatively large, with 624 pupils from 37 schools in 12 local authorities taking part. The results were complicated.

163 pupils were chosen as role models and there were no significant effects of the small group work for these pupils. Of the 461 pupils selected for extra support, there were small, but statistically significant, positive effects of small group work on at least one measure for each of the four themes studied. For New Beginnings, there was a small to medium effect (0.44) on pupils' ratings of their social and emotional skills. For Going for Goals, there was a small effect (0.37) on staff ratings of pupils' social and emotional skills, although it did not reach statistical significance, with similar results for staff ratings of children's behaviour and emotional wellbeing. There was also a small to medium effect (0.45) on pupils' ratings of their social and emotional skills. For Getting On and Falling Out, there was a limited impact on pupils' ratings, with

a small increase in social skills. For Good to Be Me, there was a decrease in teacher-rated problems during the intervention phase, however, this was smaller than reductions seen in the baseline phase. Gains were maintained at seven-week follow-up. There were no positive ratings from parents for any of the four interventions. Children in case study schools were also assessed on measures of social skills and emotional understanding but there were no significant changes. There were also some unusual findings such as a *reduction* in staff-judged empathy for Getting On and Falling Out and unexpected changes during baseline phases.

The strengths of Humphrey et al's (2008) study are the large sample size, use of a comparison group, real-life effectiveness trial format and collection of qualitative data on implementation alongside the quantitative impact study. Some difficulties with this research include the up to 57% attrition rate for parental questionnaires and not checking fidelity to the intervention guidance. This latter criticism is a particular concern, given that even one of their lead practice schools deviated significantly from the model.

Humphrey et al's report culminated in good practice recommendations: allocating enough time and space, a triangulated referral procedure, a facilitator with good rapport with the group and good modelling of skills, an appropriate setting, additional support in the classroom, fun and enjoyable activities, explicit links with wave one SEAL, fidelity to national guidance and giving small group work a high profile in school. Also, they advocated changes

such as longer or more intensive intervention, more involvement with families and standardised training at local authority and national level. However, the silver set materials and guidance have not been adapted since the publication of Humphrey et al's (2008) research to this author's knowledge.

Downey and Williams' (2009) evaluation of wave two SEAL focused on the gold set materials for families. The intervention involved workshops for parents or carers led by school-based facilitators and sessions with children and parents designed for parents to apply what they had learned in the workshops. The pilot evaluation took place with six schools in Dorset. Pre and post measures of parents' and teachers' views of children's emotional literacy were taken. For 'non-concern' children, there were no significant differences from pre to post testing for parent-rated items and a significant increase in teacher-rated self-awareness. For 'concern' children, there were significant increases in parent-rated scores for motivation and in all aspects of teacher-rated emotional literacy. Qualitative evidence after the programme revealed that parents valued the chance for quality one-to-one time with their child and social networking with other parents. They concluded that Family SEAL was most effective for those children who had been identified as causing concern in their social and emotional development and that teachers saw more of an effect than parents.

Close scrutiny of the study is difficult since it is unpublished and in the form of a conference presentation, so lacks detail. For example, there is no

information on how children and their families were selected to take part, why 'non-concern' children were included, how many people started the project and what the rate of attrition was. Downey and Williams (2009) highlight some limitations of this pilot study, such as lack of pupil voice, lack of control group, lack of information about the lasting impact and lack of information on the wider school impact. They propose a waiting list control group design with pupil measures, follow-up measures and focus groups to overcome these limitations in future.

In summary, there have only been two studies of wave two SEAL. One of these was a pilot study with no control group, which leads to uncertainty about the reasons for improvements. In both studies, parents' ratings were less favourable than teachers'.

Evaluations of wave one Primary SEAL

Hallam et al (2006) evaluated the effectiveness of the four strands of the Behaviour and Attendance pilot: continuing professional development, school improvement, SEAL curriculum materials and small group intervention. The last two are most relevant to this study. The SEAL curriculum materials were evaluated using several methods including interviews, questionnaires for staff and pre and post questionnaires for children. The report shows positive effects, for example 91% of teachers felt their confidence had increased, but this was based on a small number of returns (32 for that question). Also, data often concerned *perceptions* of change (for example in attendance) rather

than direct measures. Parental questionnaires (26 in total) were generally positive. Pre and post pupil data indicated no significant change for pupils in Key Stage 1. Key Stage 2 had some positive results but, with no control groups the authors were unable to say whether this was caused by the SEAL curriculum or maturation.

Hallam et al (2006) also evaluated groups (such as Webster-Stratton, *not* silver set groups) using interviews, a pre and post questionnaire to children, the Strengths and Difficulties Questionnaire (SDQ) with parents and teachers and case study proformas with group leaders. No control group data were collected. Positive outcomes were reported in interviews. There was a high level of attrition with the SDQ, with teacher data on 517 children pre and 145 post the intervention, for example. There were significant differences between pre and post teacher ratings on emotional and pro-social scales (although it is unclear whether the desirability of a *higher* pro-social score was appreciated). For parent ratings there were no significant SDQ differences, but positive findings from questionnaires and interviews. Pupil questionnaires showed a small negative change for social skills and relationships in Key Stage 1 and increases in social skills and perceiving own emotions for Key Stage 2.

Craig (2007) discusses many criticisms of the evaluation of the behaviour and attendance pilot. Firstly, it was published after the SEAL guidance was issued. Secondly, the results are susceptible to the Hawthorne effect (where change in behaviour is due to it being studied) due to the lack of control groups.

Another problematic factor is that participants in the qualitative evaluation (interviews and questionnaires) were not selected randomly, but came from 'good practice' schools or were recommended by local authority co-ordinators. Also, little evidence was gathered on parents' views, with parental responses from the 5000 children involved in the pilot at a rate of 0.5%. For children's views, 21 groups were interviewed, yet only one quote is used.

Craig (2007) provides a different interpretation of Hallam et al's (2006) findings. Regarding exclusions, Headteachers disagreed that SEAL had decreased fixed-term exclusions. Regarding attendance, there were no significant differences between pilot and non-pilot schools on authorised and unauthorised absences. She argues that SEAL was associated with a decrease in standards in English and maths at key stage one, and that the improvements in key stage two were due to the schools starting from a lower baseline and therefore having more scope for improvement. She suggests that decreases in children's self-reported skills are covered up by difficult to read tables. Whilst acknowledging the teacher's views on SEAL as effective in improving respect amongst pupils and reducing bullying, Craig argues that children's baseline levels do not indicate a problem to address.

The DfES (2007a) provide a case study of a school implementing SEAL, which includes information about evaluation. The intervention involved whole-school SEAL assemblies, regular SEAL lessons, high-profile visual support for SEAL, small group SEAL learning opportunities, contact with parents and

assessment for learning. The report claims an impact on attainment at the end of key stage two, improvements in attendance, attitudes to learning and self-esteem and fewer exclusions and behaviour incidents. However, no information is given on how data were collected and whether this pattern could be explained by chance or other changes in the school.

Overall, there is very little evaluation of wave one Primary SEAL given its nationwide roll-out. Neither study used a control group design, and both were rather small in scale (the Hallam et al, 2006, study was small scale for post-test data), compared with Humphrey et al's (2008) study, to justify adopting SEAL. The next section examines evaluation of Secondary SEAL.

Evaluations of Secondary SEAL

The government commissioned Smith et al (2007) to evaluate the pilot project that preceded Secondary SEAL. The pilot project involved support from local authority behaviour and attendance consultants, teaching materials, network meetings, support with developing an action plan and limited funding. The study involved 50 schools from six local authorities. Methods used to evaluate the pilot were telephone interviews with key local authority staff, visits to interview staff and pupils at ten case-study schools and questionnaires to teachers and teaching assistants in ten case study schools.

There were many aims of the project including perceptions of barriers to implementation and ideas for adapting the programme, however, this section

focuses on the part that evaluated the impact of the pilot project as this is most relevant to the present study. Impact was assessed by questionnaires to school staff. Around three quarters of respondents felt that the project had had 'considerable' or 'some' impact on children's social, emotional and behavioural skills, pupil behaviour, pupil emotional wellbeing and teaching and learning. Less impact was perceived for school attendance. Impact was also assessed through interviews with staff in local authorities and in case study schools. These people felt that the project had led to an increased awareness of social, emotional and behavioural skills, more commitment to pupil voice, a review of systems and structures, the development of a common vocabulary, implementing teaching and learning in this area, developing the role of support staff and improved collaborative working. Participants also felt that Secondary SEAL would have long-term benefits.

Although the study set out to gain information on the implementation of the pilot project rather than focusing on evaluation, the only measures of impact were those that selected teachers and local authority staff perceived, or felt might happen in future. In common with Primary SEAL, the Secondary SEAL project seems to have been extended to a nationwide programme with little objective data on its effectiveness.

2.6.3 Evaluations of other wave two social and emotional interventions

Sharp and Herrick (2000) describe anger management groups as a way to promote children's emotional literacy in schools. 45 groups were run in 38

schools in the Southampton area. The anger management groups involved six one-hour sessions led by a psychologist with a co-worker and observer from the school staff. Sharp and Herrick (2000) explain that qualitative and quantitative evaluation has occurred, but do not detail the results. Rating scales for pupils, teachers and parents were used before and after interventions, and data on exclusions and behaviour incidents were examined. They report that almost all of the pupils who participate enjoy the groups and that most would recommend the groups to friends. They claim that “the indications are that the groups do make a difference,” (page 139). The authors acknowledge the need for tighter evaluation through developing questionnaires and baseline assessments.

Maddern et al (2004) report on a multi-agency project (staffed by an educational psychologist, clinical psychologist, community psychiatric nurse and assistant psychologist) designed to promote the social skills of a group of eight boys in years five and six. The group had 20 sessions of one and a half hours duration, with weekly sessions in the autumn and spring terms and a follow-up session in the summer term. The sessions focused on making friends, thoughts and feelings and problem solving and were taught through circle time activities, games and more structured activities such as worksheets and puppet role-plays. The project was evaluated on a pre-test, post-test basis with quantitative and qualitative measures. The quantitative tests showed positive trends, with significant results for anxiety, anger management, oppositional behaviour (as rated by both teachers and parents),

hyperactivity (rated by teachers but not significant for parents) and ADHD scores. Qualitative results suggested an increase in co-operative playground behaviour and children's feelings of control over their temper. Parents meetings at school were poorly attended, but parents valued the home visits from the assistant psychologist. The authors conclude that this was a piece of successful, school-based multi-agency working. However, a criticism might be that this was quite an intensive intervention with specialist staff yet did not show a significant impact on measures of social skills and did not have a design that controlled for factors such as maturation or attention.

Coppock (2007) evaluated the effects of an emotional literacy programme on a class of Year 5 pupils. The intervention consisted of Circle Time, peer mentoring and circle of friends, although no details on sessions or materials are given. Using methods such as self-report, focus groups and peer researchers (Year 6 children from a neighbouring school who had received the intervention the previous year), she found that the Year 5 pupils had improvements in self-esteem, confidence, number of friends, relationships, helping others and learning about others' feelings. The Year 6 peer researchers also reported feeling more confident. Six parents took part in an emotional literacy programme. No information is provided on how they were selected or the content of the programme, but diaries they kept during the intervention and comments made in focus groups revealed, amongst other themes identified by the author, that they had an increased awareness of feelings, had learned to stop and think were more able to express their

feelings. This latter effect also meant that they were less likely to shout at their children, felt better for sharing feelings and had developed a sense of trust in the group. Staff who were interviewed noticed an impact on children's behaviour, emotional wellbeing and relationships. Coppock (2007) concludes that the study shows the success of the project and the value of emotional literacy work on promoting children's mental health. She argues that the action research approach was suitable as it embraced the complexity of the situation and involved young people. However, there are many unanswered questions (such as what were the results of the self-esteem assessments and how often was the intervention delivered and by whom, for example) and there is no discussion of the objectivity of the data (for instance, how was qualitative data analysed and were the interviewers or focus group facilitators involved in the delivery of the intervention?)

Parton and Manby (2009) evaluated social skills groups for Year 7 pupils in two secondary schools in the north of England. The activities were based on social skills training, role play and cognitive-behavioural techniques. Each group had up to 12 children, and participants were chosen on the basis of unwanted playground or classroom behaviour, being emotionally withdrawn or being in need of improved social skills, although it is unclear how these were identified. Pupils who were actively involved with children's social care were excluded from the study. The groups ran for 10 to 12 sessions, and each session lasted between one and two hours. They were led by two or three qualified social workers from the National Society for the Prevention of

Cruelty to Children and supported by one or two members of non-teaching school staff (for example Learning Mentors). They found that behaviour ratings changed statistically significantly for teachers' and parents' ratings but not for children's from pre to post-test. Effect sizes were small to moderate, ranging from 0.21 to 0.34 depending on which group rated. When results were analysed according to gender, the effect size was larger for girls (0.41) than boys (0.22). There were no racial differences. For self-esteem, improvements were found particularly for peer self-esteem, although no data are reported. Qualitative data indicated that children and parents noticed improvements in children's self-confidence. The children interviewed up to 18 months later still had fond memories of the groups and group leaders. Parents tended to notice more improvement than group leaders or teachers. Group leaders rated children's engagement, contribution, learning and practising new skills. Of the 38 children, 15 were rated as progressing very well, 12 as well, 10 had made some progress and three made a lower level of progress. One intention of the project was for school staff to take over responsibility for running the groups, although the authors write that this was difficult "because of the level of skill transfer involved," (page 14).

To summarise, most of the interventions were delivered in schools by professionals other than teachers. Most studies had pre and post measures of social and emotional skills and show some positive results on these, although results should be interpreted cautiously due to the small number of studies and lack of control groups.

2.6.4 Evaluations of other wave one social and emotional interventions

Matthews (2004) investigated the effects of developing children's emotional literacy through embedding it in the science curriculum for a year, rather than as a separate taught component. Pupils came from two co-educational London comprehensive schools. They were in Year 7 and had performed below average in their Year 6 SATs at primary school. 82 children took part in the collaborative group work intervention, and 83 pupils in parallel classes formed the control group. The intervention consisted of mixed-gender collaborative group activities (for example filling in a worksheet on radiation but with only one sheet per group, so consensus had to be reached) followed by completing proformas and discussions on the process to increase pupils' self-awareness. This was also supplemented with data on the group process from peer observers during the early stages of the intervention. The control group followed the same scheme of work in terms of scientific content and practical work, but without the collaborative group element. Data were collected through questionnaires (about science lessons and about feelings), interviews and opinion sheets during the intervention. Results suggested that the intervention group had more favourable attitudes towards science, developed a better understanding of opposite sex classmates, enjoyed collaborative work more and were more supportive of peers. The author concludes that bringing the affective domain into science lessons can have an effect on pupils' emotional literacy and attitudes towards science. This study might have been more persuasive to school leaders if data had also been

collected on science attainment at the end of Year 7. It might have been strengthened if the control group were also given the affective questionnaire.

Kelly et al (2004) conducted an exploratory qualitative study of the Promoting Alternative Thinking Strategies (PATHS) curriculum in a class of 9 and 10 year-olds in a primary school in Scotland. PATHS is a whole-school emotional literacy approach for primary pupils that was developed in the USA. It focuses on the same five areas as SEAL, but is more structured and is manualised. They found a high level of overall satisfaction with PATHS, using a variety of measures such as questionnaires (for pupils, teachers, the Headteacher and the Home/School worker), the Taxonomy of Problem Situations, the Kusche Affective Interview and monitoring seven target pupils. They acknowledge that the limitations of the study are its small sample size, lack of control group, impressionistic data and lack of inter-rater reliability measure.

Curtis and Norgate (2007) also evaluated PATHS. The study used a quasi-experimental approach with 114 children following the PATHS curriculum and 173 children in waiting list control groups (schools scheduled to take part in the PATHS training in future). The training for intervention group schools consisted of two days initial training by educational psychologists, which was cascaded into schools by attendees, and termly support groups. Teachers completed the SDQ as a pre and post measure and interviews were conducted with a sample of teachers from intervention schools. Pre-test scores between the intervention and control groups were significantly different ($p < 0.001$) with

children in the intervention group showing higher levels of 'problem' behaviour. The changes in scores from pre to post-test were significant for the intervention, but not the control, group. Interview data included support for and benefits of the pupil of the day, involving all staff and involving parents parts of the intervention. Regarding how PATHS had helped, themes included building feelings vocabulary, describing feelings, recognising others' feelings, empathy, managing feelings, co-operation and dealing with problems. The authors address the limitations of the study. Firstly, the control group schools had lower levels of difficult behaviour as assessed by the SDQ at pre-test, and therefore had less scope for improvement. Secondly, the teachers completing the SDQ were aware of the anticipated outcomes, which may have affected the scores they gave.

Three evaluation studies of universal emotional literacy interventions have been described. Two evaluated a specific programme, whereas the third used the science curriculum as a medium for teaching these skills. Two studies used a control group design, and the other was more exploratory. The studies found positive results for the emotional literacy interventions, but these were complicated by some methodological difficulties.

2.7 Evaluations of SEAL in Bedfordshire

Although there is relatively little published research into SEAL, as shown by the first systematic review, more informal evaluations of SEAL have occurred but remain unpublished. This section examines the evaluation of SEAL that has occurred in Bedfordshire (prior to the creation of Central Bedfordshire) at a local authority level, although it is entirely possible that evaluation has also happened at the school level.

Bedfordshire County Council and ten of its schools became involved in SEAL in 2004 by taking part in the Primary National Strategy Associate Pilot. In 2005, the SEAL materials were produced and rolled out to selected schools after training from local authority staff. During the summer term in 2006, SEAL implementation was monitored based on DfES guidance. The 40 schools that had adopted SEAL were visited by members of the school improvement, behaviour support and educational psychology teams. This involved a self-evaluation grid, discussion with the Headteacher and SEAL Co-ordinator, observation of a SEAL lesson, a walk around the school and discussion with a group of pupils. Spencer and Cuthill (2007) report on the results of this local evaluation. Common themes from the qualitative data were that schools had seen improvements in children's articulation of feelings, negotiation and writing and an impact on bullying. Curriculum materials had generally been used during PSHE time and assembly materials had been used widely. There was less use of the purple set materials (for staff) although staff meeting time had been devoted to SEAL. There was also less use of the gold set materials

(for parents). The majority of the original ten pilot schools had used the silver set materials (for small groups of pupils) and many of the other schools planned to use them. The majority of staff training and discussion about SEAL had not involved midday supervisors. In general, schools needed to consider monitoring and evaluating the impact of SEAL on pupils more.

The 2006 evaluation also involved quantitative measures for randomly selected children in the ten pilot schools, taken before SEAL and again after the first year, which Spencer and Cuthill (2007) also report on. The measures were an emotional literacy checklist for pupils and teachers (about pupils), a teacher questionnaire, SATs results, attendance figures and exclusion figures. There was a good response rate, with full data sets from seven schools and partial data from the remaining three. There was no significant improvement in children's emotional literacy and a slight downward trend, which seems alarming given the aims of SEAL but is in common with national evaluation findings. There was a significant difference in staff feeling more able to influence pupil behaviour. There were no differences in attainment, attendance or exclusion, although changes were thought likely in the long-term once SEAL was embedded. The study would have been strengthened by having comparison schools, perhaps schools that had not adopted SEAL or had chosen the Values package as an alternative, but given the practical constraints on field research this was an impressive attempt to evaluate SEAL in a local authority.

Bedfordshire County Council (2008) commissioned further monitoring of SEAL in 2008, when 103 schools from four cohorts had undertaken training. Seven schools in the local authority were selected with the aim of providing a fair representation of areas, cohorts and types of schools (urban and rural, large and small). The embedding of SEAL was assessed by a tour of the school by pupils, meeting the Head Teacher or SEAL co-ordinator and talking about SEAL and school with pupils. Each school completed a self-evaluation grid and a questionnaire. Key findings were that one school had not yet introduced SEAL, staff perceived improvements in pupils' emotional vocabulary, all schools were judged as welcoming, pupils saw school as happy and safe and said they had learned about being a good friend and anti-bullying, training was seen as a priority (particularly for lunchtime supervisors and new staff) and the need to work more closely with families was identified. It would have been interesting to have tested the hypothesis from the previous evaluation that the improvements one would hope to see with SEAL occur over the long-term, and to have taken follow-up measures from the schools that had provided pre and post data in the Spencer and Cuthill (2007) study. Also, only seven schools out of 103 were visited so it is possible that the sample was not representative of widespread practice. The fact that one of the seven schools had not yet adopted SEAL raises questions about how many other schools had not yet implemented SEAL. Other problems were that measures were largely subjective and that a development area from the first study (involving parents) was still an issue in the second study.

2.8 Gaps in the literature and rationale for the present study

Having examined the literature around SEL and SEAL, it is apparent that there has been very little research in the UK on this area. SEAL has been rolled out on a national level and yet there are only five examples of published research into it, with mixed results. It is possible that SEAL has no effects on children's social and emotional competence, or worse still is damaging. There is clearly a need for more research on the SEAL initiative.

Zins and Elias (2007) write that the evidence-base for SEL programmes is stronger than other areas of educational research. However, they argue that future research should include more replicable interventions. Similarly, Robson (2002) argues that replication is "a major cornerstone of natural science" and is necessary before a finding can be regarded as secure, yet is unfortunately rare. Therefore it would be beneficial to replicate the sole study on the silver set materials (Humphrey et al, 2008) in order to question or support its results.

Zins and Elias (2007) also argue that programme fidelity leads to better outcomes. Humphrey et al (2008) had difficulties with how schools implemented the programme and recommended programme fidelity, so there is a need for research that includes information about this.

Kimber et al (2008) distinguish between efficacy and effectiveness trials, with the former usually delivered by researchers, sometimes in specialised settings

and the latter occurring in more naturalistic, community settings. There is a need for effectiveness trials to find out the impact of interventions in the 'real world:' in other words, delivered by school staff within usual school resources.

Government-commissioned research has used measures of emotional literacy and problem behaviour. However, according to Sharp (2001), an emotionally literate person is likely to have high self-esteem. Self-esteem is a judgement of one's own self-worth and attitudes towards the self (Coopersmith, 1967). Humphrey et al (2008) suggested that there may be more generalised gains for children taking part in SEAL silver set groups, for example self-esteem. Qualter, Gardner et al (2007) suggest that some view trait EI as concerning self-perceptions of ability and ability EI as actual abilities and that this has implications for how to improve either of them. Trait EI may respond better to interventions that concern self-esteem and other self-beliefs whereas ability EI may be most influenced through direct teaching of skills. It could be useful to investigate the impact of SEAL groups on self-esteem.

Some studies have found 'sleeper effects;' positive effects that are present at follow-up but not at immediate post-test (for example Pössel et al, 2004). Although Humphrey et al (2008) took follow-up measures, these were only analysed where gains had been made at post-test, therefore there is a need for a study where all follow-up measures are compared with pre-test scores, not just where post-tests were favourable.

Another issue is locality. Zeidner et al (2002) argue that, because most of the research has been carried out in the USA, it is unclear whether SEL programmes are appropriate for children of other nationalities and cultures. Wells et al (2003) found that 15 out of 17 studies in their systematic review were from the USA and involved high proportions of children from ethnic minority groups and schools in areas of high socio-economic deprivation. Parton and Manby (2009) write that the literature evaluating social and emotional skills teaching is largely from the USA. They add that, particularly for small group work, there is a “relative paucity of published studies,” (page 7). Humphrey et al (2008) comment that “it is noteworthy that there has been comparatively little research in the UK in this area,” (page 12) leading to a lack of confidence about transferability of results.

In Tennant et al's (2007) review, a criterion for scoring reviews was applicability to an urban UK context. However, Central Bedfordshire has rural areas and unique features. For example, it is a county that has become a unitary authority with its major urban areas (Bedford and Luton) forming separate authorities. Also, there is a three tier education system (lower, middle and upper schools, rather than primary and secondary for which SEAL is designed). This means that there are many small lower schools, rather than the larger urban schools often represented in research. There has yet to be any formal evaluation of SEAL at the wave two level in the local authority, whereas wave one work has been described previously and wave three work

is likely to occur as part of the Targeted Mental Health in Schools project (Wolpert et al, ongoing, cited in Humphrey et al, 2008).

Finally, there is a need for research that uses a design that lends itself to drawing conclusions about the effectiveness of interventions. Weare and Gray (2003) write “evaluation in England is not well-developed. There is a need for much more evaluation which uses controls where appropriate and, as a minimum, before and after evaluation,” (page 6). Weare (2004) adds that that much of the evaluation work in this area has been qualitative, with very little use of pre and post designs and almost no use of control groups. This view was substantiated by the systematic review conducted for this study.

In summary, there is a need for research that: evaluates SEAL, replicates previous research, is easily replicable itself, considers programme fidelity, reflects real world practice, considers global outcomes such as self-esteem, considers longer-term effects for all participants, is carried out in the UK, reflects the locale of Central Bedfordshire and other non-urban environments and uses designs with pre and post measures and a control group. This study aims to examine these neglected areas and the next chapter explains how this was done. However, before this, the research questions are stated.

2.9 Research questions

Having reviewed the literature and considered gaps in the existing research, the following questions are posed:

- What are the effects of the targeted (wave two) Primary SEAL small group intervention (silver set materials) on social and emotional outcomes for children in lower schools in Central Bedfordshire?
 - Do pupils rate themselves as more skilled after the intervention?
 - Do parents rate pupils as more skilled after the intervention?
 - Do teachers rate pupils as more skilled after the intervention?
 - Are the results more positive immediately after the intervention or several weeks later?
- Does the targeted (wave two) Primary SEAL small group intervention (silver set materials) have an effect on children's self-esteem?
- What is the level of fidelity to the government guidance on the targeted (wave two) Primary SEAL small group intervention (silver set materials)?

CHAPTER 3: METHODOLOGY

3.1 Introduction to the methodology chapter

This chapter outlines the procedure that was used in the study, including the research design, data gathering and data analysis. However, before this there is a consideration of epistemological issues, because “researchers must first examine their underlying assumptions about the nature of reality and knowledge to make sensible decisions about all of the other steps in the research process,” (Mertens, 1998, page xiv).

3.2 Epistemological issues

The major philosophies in psychological research are arguably positivism and constructivism. These, and the debate around them, are discussed below.

3.2.1 Positivism

The positivist view holds that methods from physical science can be applied to the study of the social world. This includes the assumptions that social research can be value-free, only one reality exists for the researcher to discover (Mertens, 1998) and that the purpose of science is to establish causal laws (Robson, 2002) through hypothetico-deductive (hypothesis testing) approaches (Coolican, 2009).

Another important feature of positivism is objectivity. Positivism contends that the researcher is objective and dispassionate (Mertens, 1998). The quest for objectivity may lead to highly standardised and pre-determined procedures. A less extreme view is post-positivism, which acknowledges the limitations and influence of the researcher and therefore deals in probabilities rather than attempting to 'prove' theories.

3.2.2 Constructivism

In constructivism, reality is considered to be socially constructed. The implication for research is that there are multiple constructions of reality rather than a single truth to uncover. Also, the constructivist researcher interacts with participants, rather than being detached, and describes the context (Mertens, 1998). Other features of constructivist approaches are that procedures are often not standardised, since the goal is to discover individuals' experiences and research questions and theories can evolve during the study rather than being established at the start (Mertens, 1998). Similarly, Coolican (2009) suggests that constructivist researchers often generate, rather than test, hypotheses, which can lead to unforeseen areas of study.

3.2.3 Epistemological debate and the link with methodology

Coolican (2009) argues that some social researchers have rejected positivism. One criticism arises from the constructivist concept of multiple constructions of reality, which renders the positivist search for a single truth irrelevant.

Robson (2002) discusses philosophical criticisms of positivism, such as the rejection of the view that science should only concern that which is directly observable by the researcher and the difficulty of separating facts and values. Cohen et al (2000) write that positivism has been criticised for being reductionist (for example focusing on small parts rather than viewing people holistically and in context), dehumanising (for example denying determinism and implying passivity) and being generally inadequate for explaining individual, subjective, inner experience. Coolican (2009) adds that positivist research can be superficial or simplistic due to data gathering tools, artificial situations and a failure to examine the researcher-participant relationship.

However, the adoption of constructivism has not gone without criticism from social scientists with a more positivist stance. Coolican (2009) discusses attacks on constructivism such as the potential for greater individual bias, problems with reliability and validity and difficulties with replication, generalisation and making broad statements about social behaviour.

The two philosophies of science have been strongly associated with types of data. Positivist and post-positivist research is primarily quantitative (Mertens, 1998), whereas constructivism often gathers qualitative data in order to describe an individual's experience accurately. Cohen et al (2000) also make the link between positivism, quantitative methods and methodological issues. They argue "where one subscribes to the view which treats the social world

like the natural world... then scientific investigation... will be predominantly quantitative," (page 7).

Debate has raged about whether quantitative or qualitative data is more appropriate in social science research. Coolican (2009) suggests that quantitative methods are more objective and reliable, enabling the researcher to make comparisons across studies and to generalise to larger numbers of people. Supporters of constructivism might argue that describing human behaviour in quantitative terms is reductionist, whereas qualitative methods lead to richer and more realistic information (Coolican, 2009). However, since the data is so personal it is also less generalisable.

3.2.4 A pragmatic view

Robson (2002) suggests that the positivist versus constructivist debate is unproductive. Views can become artificially polarised; many psychologists may not entirely reject one philosophy or type of data in favour of another, but may choose the method appropriate to the task, or use a mixture of methods. Another reason that the debate is an over-simplification is that other paradigms influence psychology, for example emancipatory approaches (Mertens, 1998) and critical theory (Cohen et al, 2000). Robson (2002) prefers to discuss fixed or flexible research designs, rather than quantitative or qualitative methods.

3.2.5 Real world research

Robson (2002) argues that field research can lead to practical difficulties with random assignment to conditions, validity, ethics and control of variables for example. However, there are advantages of conducting research outside the laboratory, such as improved generalisability, decreased demand characteristics and easier access to participants (Robson, 2002).

As an applied researcher, the notion of 'real world research' is perhaps as important an influence on this research as the philosophy of science. The present study occurred in school settings with school staff delivering the interventions. Therefore the setting was naturalistic rather than being conducted in a laboratory, and carried out under realistic, rather than optimum, conditions. This study combines consideration of applied aspects of the intervention, such as contextual factors and fidelity to the guidance, with using scientific methods to answer the research questions.

3.2.6 Chosen stance

Given the subject matter of the intervention, developing children's social and emotional skills, a qualitative approach seems appropriate at first, since social and emotional skills are not easily quantifiable. However, given the explanatory (rather than explorative, descriptive or emancipatory, Robson, 2002) research questions, the political emphasis on evidence-based practice described in the introduction and the comparison with the only other study of

the small group SEAL intervention described in the literature review, a fixed design, quantitative study is likely to be more suitable.

This project adopts a primarily post-positivist and quantitative approach to exploring the research questions. This arose from taking a pragmatic view rather than because of a preference for one philosophy of science over another. However elements of the constructivist approach are retained, such as collecting information about the context of the research and a post-positivist rather than positivist stance is adopted as the limitations of the research are considered and findings are discussed in terms of probabilities.

In summary, the methodology and research design were influenced by:

- the research questions,
- previous research in this area,
- the imperative to create an evidence-base for interventions and
- the practicalities of conducting research in the 'real world.'

3.3 Research design

3.3.1 Research designs relating to the type of research questions posed

The reasons for choosing a post-positivist, fixed design have been stated. The major quantitative designs are outlined in the following sections.

Experiments

Experimental designs are considered powerful in demonstrating cause and effect relationships (Coolican, 2009). Experiments involve manipulating an independent variable, measuring dependent variables and holding other variables constant. This control of other variables reduces the likelihood of alternative explanations for the results.

One decision with experimental designs is whether to use an independent samples (between subjects) or repeated measures (within subjects) design. In the former, there is an experimental group and a separate control group. In the latter, each participant takes part in both conditions. Within subject designs are useful in eliminating the effects of participant variables since each participant is compared with themselves, however they are not suitable for all situations, for example where the effects of the intervention may be long-lasting, as the hope would be with the SEAL small group work. The effects of the intervention may persist if the child took part in the control condition second, which makes a repeated measures design unfeasible in this situation.

An alternative to the between or within subjects decision is a matched-pairs design, where participants are matched, often on the basis of pre-test scores, and then one member of the pair is randomly allocated to the experimental condition. However, this design can become very complicated when participants are matched on several variables. Another possibility is the interrupted time-series design, which involves a single group of participants

who are tested on the dependent variables several times before and after the intervention. The theory is that, if the measurements taken before the intervention are stable and they change after the intervention, then the change can be attributed to the independent variable (Mertens, 1998). However, this design is subject to threats to internal validity. Another disadvantage is the amount of time spent on testing. A similar design, although one based on single individuals rather than groups, is the single case experimental design.

For the reasons discussed above, a between-subject design is likely to be most practical. There are many of these types of experiment but one, the randomised controlled trial (RCT), has been seen as the 'gold standard' design for ascertaining whether or not something works, particularly in an era of evidence-based practice (Robson, 2002 and Fox, 2002). This could be because, with its control group and random allocation to experimental conditions, RCTs can offset threats to internal validity (confidence that the intervention caused the outcome, see section 3.3.2). Robson (2002) describes common examples of RCTs, for example the post-test only RCT (participants are randomly allocated to a treatment or control, no treatment, condition, the treatment occurs for one group and post-test scores of both groups are compared), post-test only two treatment comparison (as above but with a second treatment group rather than a no-treatment control group), pre-test post-test RCT (as the post-test only RCT but with a pre-test so that pre to post-test changes can be compared) and pre-test post-test two treatment

comparison (as pre-test post-test RCT but with a second treatment group rather than a no-treatment control group).

Based on this, it would seem that an RCT design would be most appropriate. However, Coolican (2009) provides a brief critique of the experimental method, writing that experiments are not always appropriate or ethical. This is a concern given the 'real world' nature of the present study. In this study, an RCT was not considered feasible as randomly allocating children to groups may result in an undesirable group composition. Information on group composition for the SEAL small group intervention is given in section 3.5.2.

Quasi-experiments

Quasi-experiments are a similar design to experiments but lack random allocation. As discussed, random allocation is often not possible in field research. Quasi-experimental designs overcome this problem by retaining most elements of the true experiment without random allocation. Because allocation to conditions is manipulated and not random, the term quasi-experiment is used. They are more open to threats to internal validity than if randomisation had occurred. Robson (2002) proposes considering these threats and acting to guard against them, as discussed in section 3.3.2.

Robson (2002) discusses the post-test only non-equivalent groups design (participants are allocated non-randomly to two groups, one group has a treatment and the other does not, then the two groups are tested). He

suggests that this design is unsatisfactory as an experimental design as it is impossible to say whether any differences are due to the treatment or another pre-existing difference between the groups. In a true experiment, participant variables are assumed to have been overcome through randomly allocating participants to conditions. However if random allocation does not occur, as in a quasi-experiment, other steps can be taken such as administering a pre-test, which might show that the groups were equivalent on the basis of test scores. This is known as a non-equivalent control group design. This design is affected by other threats to validity, such as regression to the mean, but is considered to be a relatively robust design where random allocation is not possible (Campbell and Stanley, 1963) as the pre-test and control group add control of some variables.

Pre-experimental designs

Where control group designs are impractical, the single group post-test only design (one group has a treatment and is then tested) is a possibility. Campbell and Stanley (1963) argue that this design has “such a total absence of control as to be of almost no scientific value,” (page 6) and hint that it is unethical to allow this design in educational research. An extension of this design is the single group pre-test post-test design (one group is tested, has a treatment and is retested). Campbell and Stanley (1963) suggest that any improvements between pre and post-test could be explained by variables other than the treatment. The problem with using ‘pre-experimental’ designs as experiments is that it is difficult to interpret the results due to numerous

threats to internal validity (see section 3.3.2). Robson (2002) advises against using such designs if possible, and they were not considered further.

3.3.2 Issues of validity in experimental designs

The validity of the experimental method is crucial in post-positivist research. Campbell and Stanley (1963) discuss internal and external validity, explaining that both are important, but that improving one may adversely affect the other.

Internal validity

Internal validity refers to confidence in the causal relationship between variables, in other words confidence in the experimental design. Campbell and Stanley (1963) identify eight possible threats to internal validity, which are summarised in table 3.1.

Name of threat	Description	Ways to control the threat
History	Changes observed are due to events that happen during the study rather than the independent variable (IV)	Control group
Maturation	Changes observed are due to biological and psychological changes in participants during the study rather than the IV	Control group
Testing	Changes observed are due to participants' experience of the pre-test rather than the IV	Control group Using post-test only design
Instrumentation	Changes observed are due to using a different instrument at different time points, rather than the IV	Using the same test at pre- and post-test
Statistical regression	Changes observed are due to extreme scores regressing towards the mean, rather than the IV	Sample from across the normal curve
Differential selection	Changes observed are due to differences between groups, rather than the IV	Random assignment to groups
Experimental mortality	Changes observed are due to participants differentially dropping out of groups, rather than the IV	Determining if people that drop out of the study are systematically different from those that do not
Selection-maturation interaction	Changes observed are due to group differences in participants' maturation, rather than the IV	Random assignment to groups

*Table 3.1: Threats to internal validity
(a summary of information presented in Mertens, 1998)*

External validity

External validity refers to the generalisability of findings. Mertens (1998) explains that this can encompass population validity (generalising results

based on sampling strategy) and ecological validity (generalising to other settings). She also describes the tension between internal and external validity, since high internal validity demands a high level of control, perhaps in a laboratory, whereas high external validity is more likely in the outside world.

Campbell and Stanley (1963) discuss four threats to external validity: the reactive effect of testing (where results cannot be generalised to those who have not been pre-tested), interaction effects (where selection bias and the experimental variable interact), reactive effects of experimental arrangements (where results cannot be generalised to different settings) and multiple treatment interference (where prior treatments continue to impact).

Mertens (1998) discusses the crucial relationship between external validity and sampling strategy in the post-positivist paradigm. Threats to external validity can occur when the participants in the study do not accurately reflect the total sampling frame. One sampling strategy is probability-based sampling, which includes simple random sampling (when every member of the sampling frame has an equal chance of being selected), stratified sampling (sampling from previously established groups, for example gender, to ensure group are adequately represented) and cluster sampling (randomly sampling at the group rather than individual level, for example classrooms). Convenience sampling is not probability based, but is commonly used in research and is based on choosing participants who are readily available; its limitations should be acknowledged by the researcher (Mertens, 1998). Other

threats to external validity are poor response rates and high attrition rates. Mertens (1998) recommends that the size and effect of both are reported.

3.3.3 Chosen research design

Consideration was given to different quantitative designs. Due to its 'gold standard' status, an RCT was considered, but judged impractical for several reasons. Firstly, randomly allocating children from the school population might not uncover suitable children for this targeted intervention. An alternative approach would be to randomly allocate children who met inclusion criteria to the experimental or control condition. However, having identified children needing support, it would be unethical to deny them the intervention (see section 3.5.4 for a discussion of ethical considerations). A waiting list control group (receiving small group SEAL later) would overcome this, as would providing the control group with an alternative intervention. For staffing reasons, the former option is likely to be more practical. However, even in these circumstances, random allocation is unlikely to be popular with school staff due to the risk of undesirable group composition.

A quasi-experimental design with a waiting list control group (non-equivalent control group design) was chosen as the research design for this study. This design is considered "well worth using" (Campbell and Stanley, 1963, page 47) in situations where random allocation is not possible. The independent variable was taking part in the intervention. The dependent variable was children's social and emotional competence. As described above, the quasi-

experimental design gives the flexibility of allowing school staff to decide on group composition, which was felt to be an important feature of conducting research in the ‘real world.’ It was believed that insisting on random allocation to groups would have led to extreme difficulty in recruiting schools to the project, would have decreased external validity and might have led to a potentially less effective intervention, which has ethical implications. The chosen design retained the control group aspect of the experiment, which is important for improving internal validity. Also, the design is similar to that used by Humphrey et al (2008), which facilitates comparison of results. The design is summarised in the figure 3.1.

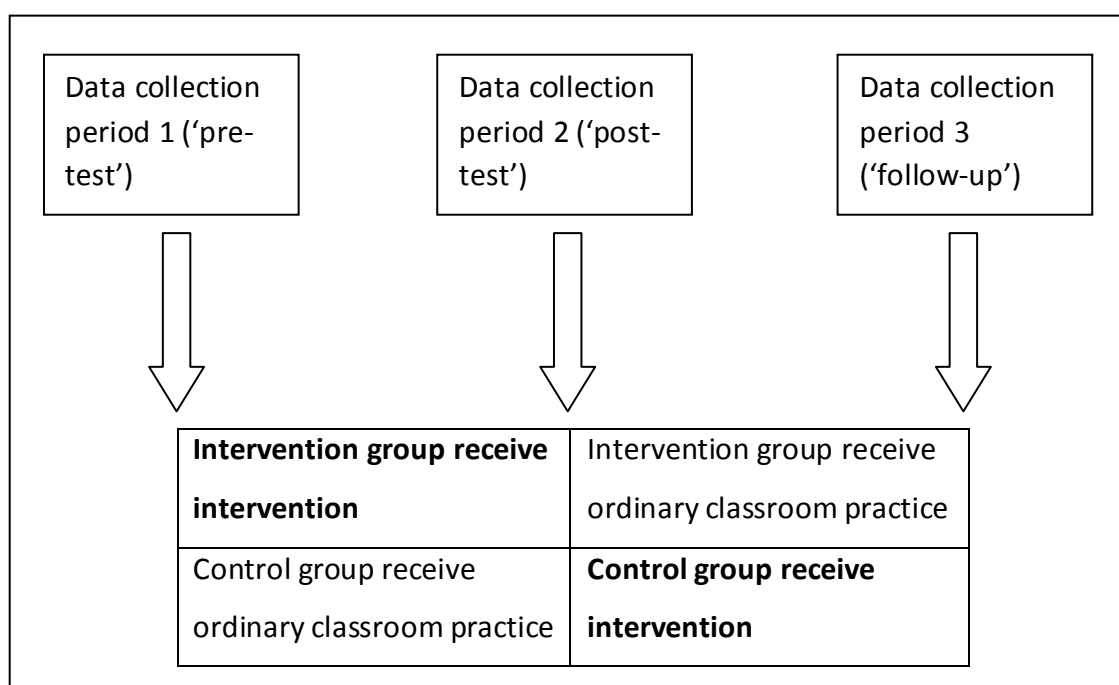


Figure 3.1: The chosen research design

3.4 Data collection

3.4.1 Problems of reliability and validity in psychological research

Validity of the experimental design has been discussed (section 3.3.2) but the validity of data collection tools is also important. The following section on reliability and validity is brief, but the concepts are revisited in section 3.4.3.

Reliability

Reliability refers to the consistency of data collection instruments. Commonly examined types of reliability are test-retest reliability (the stability of the measure at different times), internal consistency (correlations between items within the same measure) and inter-rater reliability (the correlation between two observers' ratings).

Validity

Validity concerns whether data collection tools measure what they purport to. Different types of validity include: construct validity (the extent to which the instrument measures the theoretical phenomenon; this may be informed by convergent validity which is the extent to which the measure correlates with other measures of the same construct and divergent validity which predicts a low correlation with measures of separate constructs), content validity (the extent to which the content of the measure matches the content of the construct) and criterion validity (the extent to which the measure predicts other indicators, or criteria, of the construct, for example concurrent validity

is the correlation between the measure and current behaviour and predictive validity examines the correlation between the measure and future behaviour).

3.4.2 Issues in measuring social and emotional competencies

Validity and reliability affect all researchers. However, a more specific issue in the present study was measuring social and emotional skills. There are three major approaches: self-report, informant ratings and performance measures.

Self-report usually involves a person indicating the extent to which they agree with statements in a questionnaire. Practitioners in Edmunds and Stewart-Brown's (2003) study felt strongly that pupils' views should be a key aspect of assessing their social and emotional skills. However, concerns about the self-report approach include inaccurate self-perceptions and a bias towards providing socially desirable answers (Zeidner et al, 2002). Borgers et al (2004) write that using questionnaires with children is controversial since much research into survey methodology is based on adult populations.

Using informants can involve asking other people to rate the person being assessed according to statements. Another way of using informants is to conduct a systematic observation, although a disadvantage is that information can only be gathered on behavioural, rather than internal, aspects of skills (Mayer, Caruso and Salovey, 2000a).

Mayer, Caruso and Salovey, (2000a) argue that performance measures are the preferred way to measure emotional competencies. For example, one might assess a person's ability to recognise facial expressions by labelling pictures of faces, rather than asking about that person's *belief* about their ability or another rater's *perception* of their ability. However, Mayer et al are writing from the perspective of ability EI, whereas it was argued in the literature review that SEAL is based on the trait EI, or emotional literacy, viewpoint. Another major issue with performance measures is how to determine objectively 'correct' answers. One method is to employ consensus based scoring, where the majority response is correct, but Craig (2007) writes that critics of this approach would say that this merely tests an individual's knowledge of cultural norms.

In conclusion, concerns have been raised about all three of the commonly used techniques for measuring social and emotional competence. As with the epistemology and research design, the choice of data collection methods was influenced by real world concerns; the feasibility of conducting performance measures with all participants at pre and post-test was doubted. This left a choice between self-report and informant-ratings, but both were selected as the consensus in the literature was to include a variety of respondents (see Edmunds and Stewart-Brown, 2003, and Zeidner et al, 2002).

3.4.3 Chosen outcome measures

The researcher was open to different outcome measures, and attempted to find different options through examining resources available in the local authority and searching the literature. Edmunds and Stewart-Brown (2003) reviewed instruments for measuring children's social and emotional competences, at the request of the DfES. They identified 58 instruments through literature searches and contact with academics and practitioners. This list was consulted when selecting instruments for this study.

There were several influences on the researcher when selecting instruments:

- research questions (one research question concerns the more general effects of the intervention, so a self-esteem measure was sought),
- content and construct validity (the instruments had to measure the intended focus skills of the intervention),
- suitability for the group being studied (for example age and nationality),
- fulfilling requirements of the Development and Research programme (the Strengths and Difficulties Questionnaire had to be used in the study) and
- feasibility (the researcher had a limited amount of time for data collection, and school staff had to perceive measures as feasible).

Given these influences, and that multiple outcome measures were preferable (see section 3.4.2), the following measures were selected:

Name of instrument	Intended respondent	Aspect measured
Strengths and Difficulties Questionnaire	Parent	Psychological adjustment
Emotional Literacy Assessment Instrument	Parent	Emotional literacy
	Teacher	
	Pupil	
	Teacher	
B/G Steem	Pupil	Self-esteem

Table 3.2 Outcome measures used in the study

The following sections review the reliability and validity of these measures.

Strengths and Difficulties Questionnaire

The SDQ assesses the psychological adjustment of children and adolescents (Goodman, 2001). It can be used by teachers and parents to report on children aged three to 16 and has been standardised for 11 to 16 year olds to self-report. The parent and teacher versions were used in this study. The questionnaire consists of 25 items with responses given on a three-point scale. Scores are generated for emotional symptoms, conduct problems, hyperactivity-inattention, peer problems and pro-social behaviour. An example of the version for teachers is provided in Appendix A. It is a highly favourable instrument for feasibility as it is quick to complete and freely available online. Goodman has conducted studies that show the SDQ is equal or superior to similar instruments so, given its brief nature, it may be a more practical tool for research purposes.

Goodman (1997) found good convergent validity for the SDQ, with very high correlations (ranging from 0.78 to 0.92) with the Rutter scales. Correlations between parent and teacher ratings were not significantly different for the SDQ and Rutter scales, apart from the overall score which correlated significantly higher for the parent and teacher ratings on the SDQ. The SDQ also discriminated between psychiatric and non-psychiatric populations as successfully as the Rutter scales. Overall, Goodman (1997) concludes that the SDQ is at least as valid as the Rutter scales, despite being shorter, having an identical questionnaire for parents and teachers, covering inattention and including strengths as well as difficulties. Goodman and Scott (1999) found similar results when comparing the SDQ and the Child Behavior Checklist.

Using a sample of around 10,000 British five to 15 year olds, Goodman (2001) found support for the five-factor model predicted and for three types of reliability: inter-rater, internal consistency and stability over time. The inter-rater correlations (between parents, teachers and young people) ranged from 0.21 to 0.48. Although these appear modest, Goodman (2001) argues that they compare favourably with figures reported in a previous meta-analysis (Achenbach et al, 1987, cited in Goodman, 2001), with 19 out of 21 of the SDQ inter-rater correlations higher than the average (mean) correlations reported in the meta-analysis. This suggests favourable inter-rater reliability for the SDQ compared with similar measures. The coefficients for internal consistency ranged from 0.57 to 0.88 for teacher and parent versions. The mean figure for internal consistency was 0.73, however, this was affected by

lower scores for pupil self-report (for example 0.41 for peer problems) which do not apply to this study since the self-report version was not used. Stability at four to six months after the original assessment was 0.62, which Goodman (2001) argues meets the minimum level of test-retest reliability. Again, the average was affected by lower scores for the self-report version, which was not used in this study, and the figures for teacher and parent versions of the questionnaire ranged from 0.57 to 0.82. To assess the validity of the SDQ, Goodman (2001) examined the association between the scales and psychiatric disorders. Children with the most extreme 10% of scores on the parent and teacher scales were around 15 times more likely to be at risk of psychiatric disorder.

Emotional Literacy Assessment Instrument

The Emotional Literacy Assessment Instrument (ELAI) measures self-awareness, self-regulation, motivation, empathy and social skills. It is DCSF recommended, fits the government's definition of social and emotional skills and is age-appropriate for the group studied. Faupel (2003) describes the development of the ELAI and writes that the length of time to complete the checklist was important, as well as content validity. The figures given for reliability and validity relate to a sample of 7 to 16 year olds, although the instrument used is for 7 to 11 year olds. Faupel describes the schools that took part in the standardisation procedure as a randomly selected nationally representative sample of schools in England. An example of the checklist is not provided for copyright reasons.

Reliability coefficients for the ELAI are given for the five subscales of emotional literacy as well the overall emotional literacy score for pupil, teacher and parent ratings (Faupel, 2003). A figure of 0.70 was taken as a cut off for adequate reliability. The internal consistency for overall emotional literacy for pupils, teachers and parents was 0.76, 0.94 and 0.87 respectively. When scores were broken down into the five domains, all five scored 0.70 or above for teachers and most scored over 0.70 for parents (except self-awareness and empathy which scored 0.58), whereas none of the five domains reached acceptable levels of reliability for pupils' self-ratings. For this reason, subscale scores and norms are provided for parents' and teachers', but not pupils', scores, which is why only the total emotional literacy score was used for pupils' self-ratings in this study.

Faupel (2003) examined the correlations between each questionnaire item, other items in the same subscale and the overall emotional literacy score. This revealed that items correlated most highly with other items in the same domain whilst also correlating with the overall score, indicating that subscales measured the same underlying concepts which were all related to emotional literacy rather than independent skills. Factor analyses also revealed support for the five factor model of emotional literacy. Finally, correlations between the overall emotional literacy scores for parents, teachers and pupils were calculated and were 0.29 for pupil-parent agreement, 0.42 for teacher-pupil

agreement and 0.43 for parent-teacher agreement. This indicates rather low inter-rater reliability.

Unlike the SDQ, the ELAI was also used with children in this study. Borgers et al (2004) argue that children are likely to give less reliable answers when questions are difficult, and motivation and cognitive abilities are low. They recommend four as the optimal number of response options and state that computer-assisted questionnaires are more reliable. The pupil version of the ELAI has four response options. The computerised version of the questionnaire was selected by the researcher in an attempt to increase participants' motivation. It also overcame any reading or memory difficulties by reading the questions aloud, permitting repetition and supporting response options pictorially.

B/G Steem

The questionnaire was standardised on children in British schools (Maines and Robinson, 1988). The authors aimed to use simple language structures that could be easily comprehended by children as young as six, for example by asking direct questions and avoiding double negatives. Maines and Robinson (1988) argue that being standardised on a British population and suitable for young school pupils makes the B/G Steem distinctive.

The instrument has two different versions, one for primary and one for secondary, and accordingly only the primary scale was used in this study. For

the primary sample, the test-retest correlation was 0.73 for the self-esteem items, with testing sessions occurring a week apart. No other information is provided about reliability and validity.

The questionnaires differ on one question: do you like being a boy? /do you like being a girl. As an example, the primary boys scale is in Appendix B.

3.5 Procedure

3.5.1 Timeline

Table 3.3 and figure 3.2 show the main activities carried out by the researcher in conducting this research project. Two evaluations took place, one of the New Beginnings small group SEAL intervention and one of the Getting On and Falling Out small group SEAL intervention.

3.5.2 Participants

Recruitment of participants

Sampling has been discussed in section 3.3.2 in relation to generalisability. Other influences on sampling strategy were ethics (for example consent) and practicality (for example having access to the total population).

Time period	Research activities
Autumn term 2008	<ul style="list-style-type: none"> ▪ Negotiation of the research topic with the local authority ▪ Meetings arranged with key contacts in the local authority with responsibility for SEAL
Spring term 2009	<ul style="list-style-type: none"> ▪ Design, circulation and analysis of the survey to schools ▪ Selection of measures to be used in the study
Summer term 2009	<ul style="list-style-type: none"> ▪ Recruitment of schools: telephone calls, information packs and school visits
Autumn term 2009	<ul style="list-style-type: none"> ▪ September: Pre-testing of pupils in schools 1, 2, 3 and 4 ▪ October: Post-testing of pupils in schools 1, 2, 3 and 4 ▪ November: Pre-testing of pupils in schools 5 and 6 ▪ December: Follow-up testing of pupils in schools 1, 2, 3 and 4, post-testing of pupils in schools 5 and 6 ▪ Observations of intervention sessions ▪ Collation of parent and teacher questionnaires
Spring term 2010	<ul style="list-style-type: none"> ▪ Collation of parent and teacher questionnaires ▪ Data analysis
Summer term 2010	<ul style="list-style-type: none"> ▪ Feedback to local authority and schools

Table 3.3: Main researcher activities and when they occurred

Schools

Central Bedfordshire operates a lower (Reception to Year 4), middle (Year 5 to Year 8) and upper (Year 9 upwards) school system, whereas SEAL is divided into primary and secondary phases. The middle schools are deemed secondary schools and were therefore not included in the study. Therefore, the total sampling frame was lower schools in Central Bedfordshire.

The number of Primary SEAL small group interventions operating in this population was unknown. This was not unexpected since, according to the waves of intervention model discussed in the literature review, schools should

EVALUATION 1: NEW BEGINNINGS (SCHOOLS 1, 2, 3 AND 4)

		Group activity	
		Intervention group	Control group
Time	<i>September 2009</i>	<i>Data collection ('pre-test')</i>	<i>Data collection</i>
	First autumn half term 2009	New Beginnings intervention	Ordinary classroom practice
	<i>October 2009</i>	<i>Data collection ('post-test')</i>	<i>Data collection</i>
	Second autumn half term 2009	Ordinary classroom practice	Getting On and Falling Out intervention
	<i>December 2009</i>	<i>Data collection ('follow-up')</i>	<i>Data collection</i>

EVALUATION 2: GETTING ON AND FALLING OUT (SCHOOLS 5 and 6)

		Group activity	
		Intervention group	Control group
Time	<i>November 2009</i>	<i>Data collection ('pre-test')</i>	<i>Data collection</i>
	Second autumn half term 2009	Getting On and Falling Out intervention	Ordinary classroom practice
	<i>December 2009</i>	<i>Data collection ('post-test')</i>	<i>Data collection</i>
	First spring half term 2010	Ordinary classroom practice	Going for Goals intervention

Figure 3.2: The design, timescales and interventions

provide for children who do not make adequate progress and may choose to use the silver set materials, which are freely accessible online, independently.

To establish the extent of small group SEAL interventions, the researcher designed and circulated a questionnaire to all of the schools (see Appendix C).

The results are only an indication, rather than a true guide, due to possible

response biases. For example, schools not using the intervention might have been less likely to respond as they did not want to admit this. Another factor is response rate, although attempts were made to maximise this by designing a questionnaire that was quick to complete, had a deadline for return and was followed up with a paper and electronic reminder.

Of the 96 lower schools in the local authority, 68 responded to the survey which is a response rate of 70.83%. The results were as follows:

- 63.24% of respondents indicated that they were doing whole school SEAL,
- schools that had not adopted SEAL mostly chose Values as an alternative,
- 26.47% of respondents said they were currently running small group SEAL,
- 32.35% of respondents said they will run small group SEAL in future,
- 39.71% of respondents said they might run small group SEAL in future and
- common reasons for doubt about running SEAL groups were not knowing whether the need would arise and uncertainty about future staffing levels.

After the questionnaire was sent, a multi-step procedure was followed to recruit schools to the project, see figure 3.3 for more detail on the number of schools that dropped out at each stage to leave a total of six schools taking part (described in the 'description of participants' section). The information pack referred to in the figure is in Appendix D. The researcher acknowledges that the sample was self-selecting, due to consent being needed. This biased sample severely compromises the generalisability of the results.

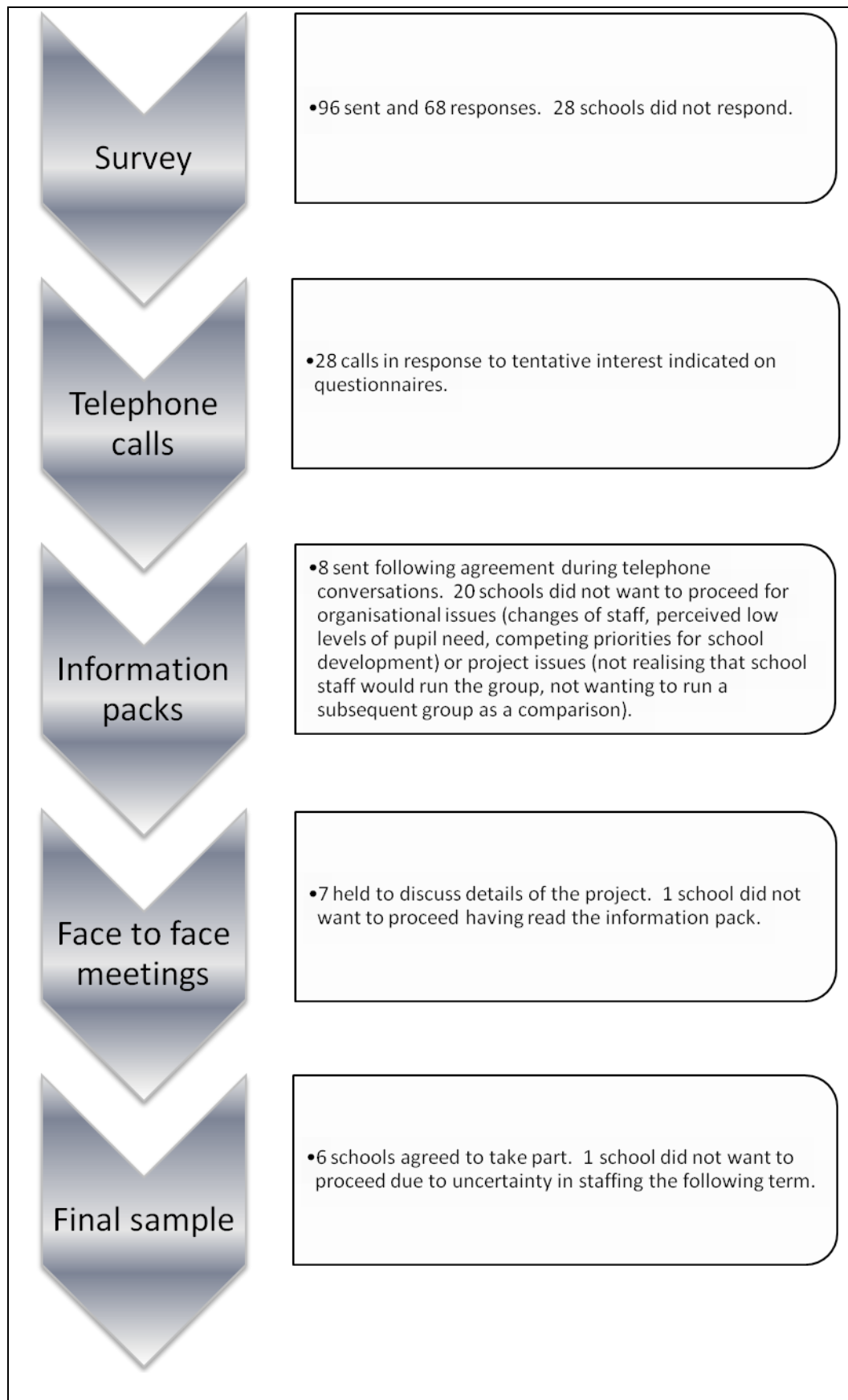


Figure 3.3 The procedure and outcomes for recruiting schools to the study

Pupils

School staff select pupils for wave two interventions due to perceived difficulties making progress. The objectivity of the selection process will vary between schools, especially as staff arguably find it more difficult to assess social and emotional development than academic achievement. Since results were compared across schools in this study, criteria for including pupils in the group were made explicit. School staff chose the pupils to participate in the groups but the researcher provided guidance in the information pack (see Appendix D). The suggested selection criteria included children with a personal plan with targets for social, emotional or behaviour issues, children highlighted by standardised assessments as having difficulties, children who were new to the school, children who had had exclusions from school or children who had faced challenging life events recently. The information pack also gave advice on group composition. A balance of need, personality and gender was recommended, along with the presence of peer role models. School staff were referred to government advice on group composition (DfES, 2006a).

Parental permission was then sought, which resulted in two children in one school and one child in another school not participating in the study (although they still took part in the intervention). Again, this introduces bias into the sample, however it was necessary for ethical reasons. Details of the pupils that took part are given in the 'description of participants' section.

Description of participants

Schools

The following descriptions of schools are based on data from several sources. One was an information gathering sheet devised by the author (see Appendix E), the next was the local authority's records of when schools attended local authority SEAL training sessions, another was the schools' most recent Ofsted report, the next was the socio-economic description of the school's postcode using the Acorn index (the local authority's preferred socio-economic indicator and replacement for free school meals data, see CACI, 2010) and the last was the author's subjective opinions formed on school visits. To protect the schools' anonymity, references for the Ofsted reports are not provided.

School 1 is a smaller than average lower school on the edge of a town in 'an area of wealthy achievers' (CACI, 2010). Nearly all pupils are White British and all speak English as their first language. The proportion of children with learning difficulties and statements of SEN is well above average. The percentage of pupils with free school meals is below average. School 1 had SEAL training in 2005/6, is a SEAL lead school and is running small group SEAL for the third year. The school and Headteacher appeared to the researcher to be supportive of the SEAL agenda, and to be keen to promote inclusive and innovative practice across the school. Examples of this are that the school hosts a county provision for children with specific language difficulties which maximises children's inclusion in their year group, rather than being a

separate 'unit,' and the SEAL groups are seen by the children as a desirable activity that they are excited to have a turn to take part in.

School 2 is an average sized lower school in a small town in 'an area of wealthy achievers' (CACI, 2010). Almost all pupils come from White British backgrounds and the proportion eligible for free school meals is below average. The percentage of pupils with learning difficulties is average. School 2 had SEAL training in 2008/9 and this is their second year of running small group SEAL. The Headteacher was happy to participate in the project as he wanted to support local authority research, although the day to day management of the school's involvement was via the Special Educational Needs Co-ordinator. This person was an experienced teacher who had managed to convey her sense of the importance of SEAL to the school's management team. For example, the Headteacher had agreed to the future delivery of SEAL groups and had been persuaded that releasing two members of staff to deliver the group would be most beneficial for the children.

School 3 is a much smaller than average primary school in a small village in 'an area of wealthy achievers' (CACI, 2010). Most of the pupils are of White British heritage. At the time of the most recent Ofsted report, no children spoke English as an additional language. The proportion of pupils with learning difficulties is average but can vary between year groups. The proportion of pupils eligible for free school meals is below average. School 3 had SEAL training in 2007/8 and this is their second year of running small

group SEAL. Although stable for the duration of the intervention, staffing at the school was somewhat unstable before and after the intervention which could be an important factor. For example the Headteacher (who agreed to the school's participation in the research) was a temporary and part-time Acting or Executive Headteacher and the group leader left the school shortly after the intervention to take a job in a middle school.

School 4 is a smaller than average school and is in a village in 'a comfortably off area' (CACI, 2010). Most of the pupils are from White British backgrounds and almost all speak English as a first language. At the time of the most recent Ofsted report, no pupils were eligible for free school meals. There are fewer than average children with learning difficulties. School 4 had SEAL training in 2007/8 and are in their second year of running small group SEAL. The Headteacher and PSHE Co-ordinator decided to participate in the project, and the author detected a sense of disempowerment from the group leaders who perhaps felt that they had been *given* the task of running the group, rather than choosing to do it based on interest and confidence. The school showed some commitment to supporting the staff by having two group leaders but this could usefully have been supplemented by supervision from a more senior and experienced member of staff perhaps.

School 5 is a similar size to most other primary schools and is in a town in 'a comfortably off area' (CACI, 2010). The majority of children are from White British families, the proportion from minority ethnic families is below the

national average and very few pupils speak English as an additional language. Attainment on entry is below average in communication, language and literacy, mathematics and personal, social and emotional development. School 5 had SEAL training in 2005/6 and this is their first year of running small group SEAL. The school (along with a cluster of local schools) had recently recruited a Family Worker and part of her role was to deliver the SEAL small group intervention. The Headteacher was perhaps partly motivated to participate by the chance to have advice from the educational psychology service on best practice in the delivery of the intervention.

School 6 is an average sized primary school in a town in 'a comfortably off area' (CACI, 2010). The proportion of pupils with learning difficulties is average, although the number of pupils with a statement of special educational needs is above average. The proportion of pupils eligible for free school meals is below average. Most pupils are from White British backgrounds and no pupils are at an early stage of learning English. School 6 had SEAL training in 2006/7, is a SEAL lead school and is in the third year of running small group SEAL. Although the school had been providing the small group SEAL intervention for some time, the group leader was new to the role and was the same Family Worker mentioned in the description of School 5. The Headteacher agreed to the school's participation in the project, perhaps with a sense of duty towards the local authority due to the school's lead practice status.

Group leaders

Seven group leaders were involved in this study. All were female. Five of the group leaders had had some training in running small group work. This came from the Behaviour Support Team and the Educational Psychology Service (who trained staff specifically in SEAL small group work) and from the local Pyramid club (more general training). The other two group leaders worked in the same school and did not have any training. In four of the schools there was one group leader in each school, in the other two schools there were two leaders each. One group leader was a teacher and Special Educational Needs Co-ordinator, one was a family worker, one was a special needs nursery nurse and four were teaching assistants. The researcher acknowledges the range of knowledge, skills, confidence and experience within this group of participants and (although this was not specifically examined in the present study) the potential for this to affect the outcome of the intervention.

Pupils

74 children took part in the study; 51 in the New Beginnings evaluation (from schools 1, 2, 3 and 4) and 23 in the Getting On and Falling Out evaluation (from schools 5 and 6). Tables 3.4 to 3.10 give demographic information about the pupils that participated in the study.

<i>Name of intervention</i>	<i>Intervention group</i>	<i>Comparison group</i>	<i>Total</i>
New Beginnings	28	23	51
Getting On and Falling Out	12	11	23
<i>Total</i>	<i>40</i>	<i>34</i>	<i>74</i>

Table 3.4: The number of children in the intervention and comparison groups

<i>Name of intervention</i>	<i>Year 3</i>	<i>Year 4</i>	<i>Total</i>
New Beginnings	31	20	51
Getting On and Falling Out	19	4	23
<i>Total</i>	50	24	74

Table 3.5: The number of children from each year group participating in the study

<i>Name of intervention</i>	<i>Male</i>	<i>Female</i>	<i>Total</i>
New Beginnings	25	26	51
Getting On and Falling Out	11	12	23
<i>Total</i>	36	38	74

Table 3.6: The number of boys and girls participating in the study

<i>Name of intervention</i>	<i>White British</i>	<i>Black (Black African, Black Caribbean, Black Other)</i>	<i>Mixed (Mixed White and Black Caribbean, Mixed Other)</i>	<i>Refused to answer</i>	<i>Total</i>
New Beginnings	48	2	1	0	51
Getting On and Falling Out	17	2	3	1	23
<i>Total</i>	65	4	4	1	74

Table 3.7: The number of children from different ethnic groups participating in the study

<i>Name of intervention</i>	<i>English is first language</i>	<i>English is additional language</i>	<i>Refused to answer</i>	<i>Total</i>
New Beginnings	51	0	0	51
Getting On and Falling Out	22	0	1	23
<i>Total</i>	73	0	1	74

Table 3.8: The number of children speaking English as an additional language participating in the study

<i>Name of intervention</i>	<i>None identified</i>	<i>School Action</i>	<i>School Action Plus</i>	<i>Statement</i>	<i>Total</i>
New Beginnings	38	9	1	3	51
Getting On and Falling Out	15	6	2	0	23
<i>Total</i>	53	15	3	3	74

Table 3.9: The number of children with special educational needs participating in the study

<i>Name of intervention</i>	<i>Targeted</i>	<i>Role model</i>	<i>Total</i>
New Beginnings	40	11	51
Getting On and Falling Out	16	7	23
<i>Total</i>	56	18	74

Table 3.10: The reasons for including the children in the interventions

Response rates and attrition

The data in tables 3.4 to 3.10 relates to pupils at the start of the study. Parents were reminded of their right to withdraw themselves or their children from the study at each data collection point. No parents requested that their children were withdrawn from the study.

For the New Beginnings intervention, a pupil from School 1 moved out of the area during the study, meaning that no data were collected for this child at follow-up. Also, in School 4, the composition of the waiting list comparison group was changed by the school without the author's knowledge. It was decided that, since there were three children who had been part of the data collection process but who were not going to benefit from the intervention, data would not be collected from them at follow-up. Accordingly, response

rates for the pupil measures (where the researcher collected the data with each child) were 100% at pre-test and post-test and 92% at follow-up.

The response rates for parents were 73% at pre-test, 54% at post-test and 55% at follow-up. The response rates for teachers were 70% at pre-test, 70% at post-test and 67% at follow-up. There was a problem at School 4, who retained the parent and teacher questionnaires for their own information but then mislaid them in a school office move over the Christmas holidays. It was decided to keep School 4 in the study since near complete sets of pupil data had been collected. When School 4 is excluded from the analysis of parent and teacher questionnaires, the response rate improves to 90%, 67% and 75% respectively for parents and 91%, 91% and 87% respectively for teachers. The teacher response rate was not expected to reach 100%, since it had been agreed with School 3 that the teacher would only complete the ELAI and not the SDQ, as she had taken on a temporary post of Acting Headteacher and so faced an increased workload.

For the Getting On and Falling Out intervention, response rates for the pupil measures were 100% at pre and post test. Response rates for the teacher questionnaires were 100% at pre-test and 87% at post-test. Response rates for the parent questionnaires were 57% at pre-test and 39% at post-test.

The implications of the response rates are dealt with in the discussion chapter.

3.5.3 Intervention

Materials

The silver set materials are freely available to school staff online (see DfES, 2006a for the booklet). They provide a structure for a typical group work session, suggested activities and assistance with planning the sessions. The materials are intended to be used flexibly and there is no script to follow, as there is in other similar interventions. Each session has a warm-up, core and relaxation activity and the booklet offers options to choose from for each of these. School staff are advised to have a predictable routine to the sessions. Additional resources are often required (for example, a digital camera, worksheets and fabric) and these are also listed in the booklet.

Frequency and duration

Sessions should occur weekly and typically be at least 40 minutes long (DfES, 2006a). An intervention usually involves six or seven sessions in a half term.

Location

Guidance is given on where to hold the sessions (DfES, 2006a). A room that is comfortable, is free from interruptions, permits circle, paired and individual activities, is attractive and personalised and provides easy access to resources is advocated.

Intervention integrity

Intervention integrity means that the intervention is implemented in the manner intended. It cannot be assumed and can prove a challenge to researchers (Mertens and McLaughlin, 2004). In this study, data were collected on the running of the intervention in the schools. Observations of each intervention in each school were carried out, along with a brief interview with the group leader, using a checklist (see Appendix F). The checklist was somewhat difficult to compile, since the materials are intended to be used flexibly. However, the suggested structure and organisational arrangements from the guidance document (DfES, 2006a) were used as the basis for the checklist. The results of the investigation into intervention integrity are summarised in the next chapter. The implications are dealt with in the discussion.

Comparison group intervention

The comparison group comprised a waiting list control condition. This meant that the children in the comparison group were due to participate in the SEAL silver set small group work the next half term. Their intervention was slightly different, as the SEAL theme changes every half term in schools. The comparison group for the New Beginnings evaluation were waiting for the Getting On and Falling Out intervention. The comparison group for the Getting On and Falling Out evaluation were waiting for the Going for Goals intervention. This information was also summarised in figure 3.2. It should

also be noted that the children in both the 'experimental' and comparison conditions were exposed to wave one SEAL during the data collection period.

3.5.4 Ethical considerations

The British Psychological Society (BPS) specifically addresses issues in research with human participants (BPS, 2008). Ethical principles are now considered using headings from these guidelines.

Consent

The group interventions were part of normal school practice, so usual school rules about consent from participants and parents (or those acting *in loco parentis*) were followed. In practice, this entailed parents completing a reply slip to indicate their agreement to their child taking part in the group work.

However, the questionnaires were additional to usual school procedures. Schools sent letters to parents seeking their agreement to complete the questionnaires, and for the researcher to complete questionnaires with their child. Parents could show their lack of consent by indicating this on their reply slip or by not replying to the letter. Parents were also assured that lack of consent to take part in the research study did not affect their child's right to take part in the group work, if they had agreed to this.

Teachers were asked for their consent to take part by completing questionnaires. They could show their lack of consent verbally or by not

completing the questionnaires. Children's verbal consent to take part in the groups was gained by school staff and their consent to take part in the project was obtained verbally by the researcher during the administration of the questionnaires, whilst also recognising that "avoidance of the testing situation may be taken as evidence of failure to consent" (BPS, 2008).

Deception

There was no deception of participants as part of the research project.

Debriefing

Schools will be provided with written feedback on the findings and offered a meeting to discuss these. School staff may wish to follow their usual procedures for sharing this information with pupils, parents and governors.

Withdrawal

Participation in the study was voluntary and participants were reminded of their right to withdraw at any time, including retrospectively.

Confidentiality

BPS (2008) guidance states that data "obtained about a participant during an investigation is confidential unless otherwise agreed in advance." The identities of individual children, group leaders, parents and schools have been protected as data has been aggregated or anonymised.

Protection of participants

The BPS (2008) advises that “risk of harm must be no greater than in ordinary life.” In this study, the interventions would have occurred regardless of data collection. Participants were not anticipated to experience more stress due to the study, but one possibility was that school staff may have worried about collecting questionnaire responses from parents. The researcher reassured them that no extra effort was required to ‘chase’ returns.

Observational research

Consent to observe the group was obtained directly from the group leader. Explicit consent was not sought from pupils as school is a place where “those observed would expect to be observed by strangers,” (BPS, 2008).

Giving advice

Unknown psychological difficulties were unlikely to be uncovered by the research, since children with existing difficulties were targeted. However, if evidence of unknown needs had been uncovered, the researcher would have informed school staff and advised on local processes for meeting these needs.

Colleagues

The researcher acted alone in most research activities. In the unlikely event that groups were run unethically and participants were at risk of harm, the researcher would have encouraged the group leader to re-evaluate their practice, in accordance with the guidance.

3.6 Data analysis

3.6.1 Descriptive

Certain data (for example mean scores and standard deviations) were described before being subjected to further analysis. Also, the research question on the fidelity of the intervention was addressed in a descriptive manner. Other important descriptive statistics include the distribution of the data. This is important as it affects the choice of test used for significance testing (parametric tests assume that data is normally distributed).

3.6.2 Significance testing

The first step in interpreting the results of a pre-test post-test non-equivalent group design is to establish pre-test differences between the two groups (Robson, 2002). The ideal is for scores to be similar, indicating that the groups were roughly equivalent to begin with. Post-test scores can then be compared, controlling for any pre-test differences if necessary.

The next step is to use statistical tests to infer whether the null hypothesis can be rejected. Coolican (2009, page 334) discusses some common misconceptions about null hypotheses. The first misconception is that phrases such as 'there will be no significant differences between experimental and control group scores' can be used to phrase a null hypothesis. This is because the null hypothesis is not a prediction but rather is a claim about the population. Also, some difference would be expected due to sampling error.

Secondly, it is untrue that the probability that the null hypothesis is true is calculated, rather it is the probability that the results would occur if the null hypothesis was true. This probability is the p result given in inferential statistics. Conventionally, the cut-off point for rejecting the null hypothesis is when the probability of the effect occurring by chance is equal to or less than 0.05, and this convention is adopted in the present study. For the purposes of clarity, the research questions are now restated as null hypotheses:

- If the null hypothesis is true, there would be little difference between intervention group and comparison group scores on measures of social and emotional outcomes
- If the null hypothesis is true, there would be little difference between the intervention group and comparison group on the self-esteem measure

Inferential statistics were used to test the probability of the results occurring if the null hypothesis is true. If the p value was greater than 0.05, there was not enough evidence to reject the null hypothesis. If the p value was 0.05 or less, the null hypothesis was rejected, meaning the results may be explained by an alternative hypothesis. The alternative hypotheses in this study were:

- Pupils rate themselves more favourably on social and emotional measures as a result of the small group SEAL intervention
- Parents rate their children more favourably on social and emotional measures as a result of the small group SEAL intervention

- Teachers rate pupils more favourably on social and emotional measures as a result of the small group SEAL intervention
- The results are more favourable several weeks after a small group SEAL intervention than immediately afterwards
- Pupils rate themselves more favourably on self-esteem measures as a result of the small group SEAL intervention

The use of significance testing in psychological research is commonplace but controversial (Cohen, 1994). Robson (2002) concludes that significance testing is often expected, but should not be solely relied upon. Providing information on the direction and size of the effect is also advised.

3.6.3 Effect size

Effect size is a statistic that is desirable to quote alongside significance as it is not affected by sample size (Robson, 2002). Because of this, effect size calculations were also performed on the pre to post-test score changes that were found to be significant. Effect size is usually computed using partial eta squared (the proportion of variance explained by the independent variable) or Cohen's *d* (the difference between groups in standard deviations, Pallant, 2007). In order to facilitate comparison with previous research (Humphrey et al, 2008), Cohen's *d* was chosen as the effect size statistic. However, since this relies on the mean and standard deviation, and the data were not normally distributed, partial eta squared was also used.

CHAPTER 4: RESULTS

4.1 Introduction to the results chapter

This chapter presents the findings of the study and a limited commentary on these results. For a more extensive consideration, please refer to the discussion chapter. Different levels of data analysis are provided, including descriptive and inferential statistics. As mentioned previously, the researcher used several outcome measures, each with sub-scales, resulting in a rather large data-set. To facilitate the reader's understanding of this chapter, a reminder of the 26 dependent variables in the study is provided in Table 4.1.

Name of instrument	Sub-scale scores generated			Desired direction of change
	Parent	Teacher	Pupil	
Emotional Literacy Assessment Instrument (ELAI)	Self-awareness	Self-awareness	-----	An increase in scores
	Self-regulation	Self-regulation	-----	
	Motivation	Motivation	-----	
	Empathy	Empathy	-----	
	Social skills	Social skills	-----	
	Total emotional literacy	Total emotional literacy	Total emotional literacy	
Strengths and Difficulties Questionnaire (SDQ)	Emotional distress	Emotional distress	-----	A decrease in scores
	Behaviour problems	Behaviour problems	-----	
	Hyperactivity	Hyperactivity	-----	
	Peer problems	Peer problems	-----	
	Pro-social behaviour	Pro-social behaviour	-----	An increase in scores
	Total difficulties	Total difficulties	-----	A decrease in scores
BG Steem	-----	-----	Self-esteem	An increase in scores

Table 4.1 The sub-scales and informants for each instrument used in the study

Table 4.1 highlights the large number of dependent variables in the study, which brings with it the potential for false positive results caused by chance rather than the intervention (type 1 errors). The researcher acknowledges this problem but chose not to perform a Bonferroni correction on this occasion, although it was tempting to do so given its inherently cautious preference for type 2 (false negatives) over type 1 errors. However, Bonferroni corrections can lead to an irrelevant focus (on the number of tests performed and the universal null hypothesis) and practical issues about which tests to include, as Perneger (1998) discusses. Perneger (1998) suggests that “simply describing what was done and why, and discussing the possible interpretations of each result” is a reasonable alternative way to reach conclusions without needing to perform a Bonferroni adjustment.

Another reminder for the reader, in figure 4.1, is the research design for the two evaluation studies that were carried out.

EVALUATION 1: NEW BEGINNINGS (SCHOOLS 1, 2, 3 AND 4)

		Group activity	
		Intervention group	Control group
Time	<i>September 2009</i>	<i>Data collection ('pre-test')</i>	<i>Data collection</i>
	First autumn half term 2009	New Beginnings intervention	Ordinary classroom practice
	<i>October 2009</i>	<i>Data collection ('post-test')</i>	<i>Data collection</i>
	Second autumn half term 2009	Ordinary classroom practice	Getting On and Falling Out intervention
	<i>December 2009</i>	<i>Data collection ('follow-up')</i>	<i>Data collection</i>

EVALUATION 2: GETTING ON AND FALLING OUT (SCHOOLS 5 and 6)

		Group activity	
		Intervention group	Control group
Time	<i>November 2009</i>	<i>Data collection ('pre-test')</i>	<i>Data collection</i>
	Second autumn half term 2009	Getting On and Falling Out intervention	Ordinary classroom practice
	<i>December 2009</i>	<i>Data collection ('post-test')</i>	<i>Data collection</i>
	First spring half term 2010	Ordinary classroom practice	Going for Goals intervention

Figure 4.1: The design, timescales and interventions

4.2 Descriptive statistics

4.2.1 Distribution of the data

It is important to establish whether the data are normally distributed, as this affects which tests can be used for the inferential statistics. Pallant (2007)

suggests assessing normality using the Kolmogorov-Smirnov test and by examining figures for skewness.

A non-significant result on the Kolmogorov-Smirnov test indicates normality (Pallant, 2007). Dancey and Reidy (2007) explain that a skewness value of 0 indicates that data is not skewed, whereas a positive value indicates a positively skewed distribution (scores clustered at the low end) and a negative value indicates a negatively skewed distribution (scores clustered at the high end). They argue that skewness values of 1 or -1 mean that the data are not normally distributed enough to use parametric tests.

New Beginnings

As shown in table 4.2, out of the 26 dependent variables, 19 were not normally distributed on the Kolmogorov-Smirnov test. Data for 6 variables were positively skewed and 1 was negatively skewed at above the 1 or -1 level.

Getting On and Falling Out

As shown in table 4.3, out of the 26 dependent variables, 11 were not normally distributed on the Kolmogorov-Smirnov test. Data for 3 variables were positively skewed at above the 1 or -1 level. It is important to note that the parent measures were based on a small number of questionnaires, due to sample size and response rate.

Overall summary and implications

The result of this analysis is that much of the data is not normally distributed. This is not surprising given that the SEAL small group interventions are aimed at children at extreme ends of the normal distribution curve (targeted pupils from one end and role model pupils from the other). One of the assumptions of parametric tests is that the data are normally distributed. Another assumption is that a random sample of the population has been used. Since these assumptions have been violated, non-parametric tests were used for the inferential statistics.

4.2.2 Measures of central tendency and variability

The mean is the most commonly used measure of central tendency in psychological research (Dancey and Reidy, 2007) however, given the findings on the distribution of the data, the median is also stated as it is less influenced by extreme scores. The range and standard deviation are given to show the spread of the data. The mean (M), median (Md), standard deviation (SD) and range (R) for each dependent variable and each data collection period are in tables 4.4 to 4.8. The results for the New Beginnings intervention are shown first, followed by the results for the Getting On and Falling Out intervention.

Dependent variables	Kolmogorov-Smirnov			Skewness	
	Statistic	Degrees of freedom	Significance	Statistic	Standard Error
Pupil emotional literacy total score	.106	51	.200	-.478	.333
Parent emotional literacy self-awareness score	.110	39	.200	-.388	.378
Parent emotional literacy self-regulation score	.156	39	.018	.110	.378
Parent emotional literacy motivation score	.125	39	.125	-.250	.378
Parent emotional literacy empathy score	.141	39	.048	-.133	.378
Parent emotional literacy social skills score	.210	39	.000	-1.469	.378
Parent emotional literacy total score	.116	39	.200	-.417	.378
Teacher emotional literacy self-awareness score	.183	37	.003	-.463	.388
Teacher emotional literacy self-regulation score	.157	37	.021	-.592	.388
Teacher emotional literacy motivation score	.104	37	.200	-.152	.388
Teacher emotional literacy empathy score	.157	37	.022	-.612	.388
Teacher emotional literacy social skills score	.206	37	.000	-.918	.388
Teacher emotional literacy total score	.107	37	.200	-.738	.388
Pupil BG Steem score	.207	51	.000	-.502	.333
Parent SDQ emotional difficulties score	.194	35	.002	.725	.398
Parent SDQ behaviour difficulties score	.212	35	.000	1.563	.398
Parent SDQ hyperactivity score	.238	35	.000	.913	.398
Parent SDQ peer problems score	.257	35	.000	1.488	.398
Parent SDQ pro-social behaviour score	.269	35	.000	-.945	.398
Parent SDQ total difficulties score	.167	35	.015	1.220	.398
Teacher SDQ emotional difficulties score	.143	32	.096	.672	.414
Teacher SDQ behaviour difficulties score	.346	32	.000	2.045	.414
Teacher SDQ hyperactivity score	.214	32	.001	1.020	.414
Teacher SDQ peer problems score	.318	32	.000	1.445	.414
Teacher SDQ pro-social behaviour score	.185	32	.007	-.451	.414
Teacher SDQ total difficulties score	.179	32	.010	.775	.414

Table 4.2: The results of the tests of the distribution of the data for the New Beginnings evaluation with non-normal results highlighted in bold

Dependent variables	Kolmogorov-Smirnov			Skewness	
	Statistic	Degrees of freedom	Significance	Statistic	Standard Error
Pupil emotional literacy total score	.152	23	.180	.400	.481
Parent emotional literacy self-awareness score	.178	12	.200	-.163	.637
Parent emotional literacy self-regulation score	.124	12	.200	.022	.637
Parent emotional literacy motivation score	.214	12	.136	-.420	.637
Parent emotional literacy empathy score	.238	12	.058	-.813	.637
Parent emotional literacy social skills score	.193	12	.200	-.498	.637
Parent emotional literacy total score	.145	12	.200	-.250	.637
Teacher emotional literacy self-awareness score	.162	23	.123	-.363	.481
Teacher emotional literacy self-regulation score	.204	23	.014	-.335	.481
Teacher emotional literacy motivation score	.137	23	.200	.141	.481
Teacher emotional literacy empathy score	.220	23	.005	-.427	.481
Teacher emotional literacy social skills score	.221	23	.005	-.033	.481
Teacher emotional literacy total score	.163	23	.117	-.050	.481
Pupil BG Steem score	.202	23	.016	.149	.481
Parent SDQ emotional difficulties score	.207	14	.105	.449	.597
Parent SDQ behaviour difficulties score	.218	14	.069	1.098	.597
Parent SDQ hyperactivity score	.143	14	.200	.394	.597
Parent SDQ peer problems score	.240	14	.028	1.245	.597
Parent SDQ pro-social behaviour score	.237	14	.033	-.899	.597
Parent SDQ total difficulties score	.147	14	.200	.125	.597
Teacher SDQ emotional difficulties score	.280	23	.000	.935	.481
Teacher SDQ behaviour difficulties score	.284	23	.000	.925	.481
Teacher SDQ hyperactivity score	.224	23	.004	.420	.481
Teacher SDQ peer problems score	.235	23	.002	2.024	.481
Teacher SDQ pro-social behaviour score	.230	23	.003	-.282	.481
Teacher SDQ total difficulties score	.135	23	.200	.784	.481

Table 4.3: The results of the tests of the distribution of the data for the Getting On and Falling Out evaluation with non-normal results highlighted in bold

Dependent variable	Intervention group				Control group			
	M	Md	SD	R	M	Md	SD	R
Pupil emotional literacy total score	85.33	85	8.989	26	79.75	79.5	11.795	38
Parent emotional literacy self-awareness score	12.87	13	2.2	8	13.08	13	2.503	9
Parent emotional literacy self-regulation score	12.6	12	3.641	12	12.5	12	3.233	10
Parent emotional literacy motivation score	12.07	13	3.845	12	12.92	13.5	3.147	11
Parent emotional literacy empathy score	15.6	16	2.324	8	16.17	16	2.517	8
Parent emotional literacy social skills score	17.93	19	2.492	8	16.75	17	3.166	11
Parent emotional literacy total score	71.07	72	10.767	40	71.42	72.5	11.759	40
Teacher emotional literacy self-awareness score	11.47	12	1.922	7	11.58	12	2.151	7
Teacher emotional literacy self-regulation score	12.73	13	3.24	9	11.75	12	3.251	10
Teacher emotional literacy motivation score	10.27	11	3.674	11	10.42	11	2.746	9
Teacher emotional literacy empathy score	12.6	13	2.261	7	12.83	13.5	2.855	9
Teacher emotional literacy social skills score	14.07	16	2.52	7	13.58	14	1.881	6
Teacher emotional literacy total score	60.47	59	8.692	27	63.08	66	12.788	38
Pupil BG Steem score	16.2	16	2.396	6	16.75	17	1.658	6
Parent SDQ emotional difficulties score	1.93	2	1.71	4	3	2	2.663	7
Parent SDQ behaviour difficulties score	1.47	2	1.125	3	2.33	2	2.387	8
Parent SDQ hyperactivity score	4.07	4	2.789	10	3.5	3	2.876	10
Parent SDQ peer problems score	1.67	1	1.952	7	1.42	1	1.832	6
Parent SDQ pro-social behaviour score	8.07	9	1.792	5	8.25	9	2.137	6
Parent SDQ total difficulties score	9.13	7	6.357	23	10.25	8	6.943	25
Teacher SDQ emotional difficulties score	2.8	3	2.145	7	2.42	2.5	2.314	8
Teacher SDQ behaviour difficulties score	0.73	0	1.438	5	1.08	0	2.109	7
Teacher SDQ hyperactivity score	3.07	2	3.15	10	2.83	1.5	3.538	10
Teacher SDQ peer problems score	1.8	1	2.336	7	1.17	0	1.642	4
Teacher SDQ pro-social behaviour score	7.93	8	1.831	5	7.92	9	2.466	6
Teacher SDQ total difficulties score	8.4	6	5.604	18	7.5	3.5	7.646	22

Table 4.4 The mean, median, standard deviation and range of scores for the New Beginnings evaluation at pre-test

Dependent variable	Intervention group				Control group			
	M	Md	SD	R	M	Md	SD	R
Pupil emotional literacy total score	88.54	88.00	7.644	23	87.40	87.00	6.899	20
Parent emotional literacy self-awareness score	13.15	13.00	3.184	11	13.20	13.50	2.700	9
Parent emotional literacy self-regulation score	12.92	14.00	3.883	14	13.90	14.00	2.644	9
Parent emotional literacy motivation score	12.62	12.00	4.610	14	13.80	14.00	2.781	10
Parent emotional literacy empathy score	15.77	16.00	2.555	9	16.00	16.00	2.906	8
Parent emotional literacy social skills score	18.38	19.00	2.063	7	17.50	19.00	3.689	12
Parent emotional literacy total score	72.85	74.00	13.594	46	74.40	77.50	12.094	42
Teacher emotional literacy self-awareness score	13.23	13.00	1.964	6	13.70	14.50	2.627	8
Teacher emotional literacy self-regulation score	11.85	13.00	3.236	10	12.70	14.00	3.129	10
Teacher emotional literacy motivation score	11.38	11.00	2.815	9	12.20	12.50	3.327	9
Teacher emotional literacy empathy score	13.31	13.00	2.562	8	14.60	15.50	1.955	6
Teacher emotional literacy social skills score	14.46	16.00	1.941	5	15.00	16.00	2.211	7
Teacher emotional literacy total score	64.23	63.00	10.010	27	68.20	69.50	11.840	39
Pupil BG Steem score	17.69	18.00	1.377	4	17.70	17.50	1.494	4
Parent SDQ emotional difficulties score	1.54	1.00	1.808	5	3.80	3.00	2.348	8
Parent SDQ behaviour difficulties score	2.00	2.00	1.826	6	1.30	1.00	2.111	7
Parent SDQ hyperactivity score	4.46	4.00	3.017	10	2.70	2.00	1.829	6
Parent SDQ peer problems score	1.38	1.00	1.895	7	1.70	1.00	2.710	9
Parent SDQ pro-social behaviour score	8.54	9.00	1.450	5	8.60	9.50	2.011	5
Parent SDQ total difficulties score	9.38	6.00	7.433	24	9.40	7.50	6.484	23
Teacher SDQ emotional difficulties score	1.31	1.00	1.109	3	1.20	.00	1.989	6
Teacher SDQ behaviour difficulties score	.77	.00	1.235	4	1.00	.00	2.211	7
Teacher SDQ hyperactivity score	2.08	1.00	2.660	8	2.50	1.50	2.953	7
Teacher SDQ peer problems score	1.23	.00	1.833	6	.60	.00	1.350	4
Teacher SDQ pro-social behaviour score	8.54	9.00	1.664	5	8.70	9.50	1.889	6
Teacher SDQ total difficulties score	5.38	4.00	5.059	15	5.30	3.00	6.255	18

Table 4.5 The mean, median, standard deviation and range of scores for the New Beginnings evaluation at post-test

Dependent variable	Intervention group				Control group			
	M	Md	SD	R	M	Md	SD	R
Pupil emotional literacy total score	88.36	86.50	7.712	24	83.73	86.00	11.568	42
Parent emotional literacy self-awareness score	13.36	13.00	2.678	9	12.73	13.00	1.618	6
Parent emotional literacy self-regulation score	12.86	12.50	3.231	11	13.73	13.00	3.069	10
Parent emotional literacy motivation score	12.36	12.00	4.069	12	12.82	14.00	2.523	8
Parent emotional literacy empathy score	15.86	16.00	2.568	8	15.82	16.00	2.483	7
Parent emotional literacy social skills score	17.86	18.50	2.381	7	17.00	18.00	3.742	13
Parent emotional literacy total score	72.29	72.00	12.118	39	72.09	72.00	10.931	38
Teacher emotional literacy self-awareness score	12.00	13.00	1.961	6	13.36	14.00	2.767	7
Teacher emotional literacy self-regulation score	12.50	13.50	3.391	9	13.45	16.00	3.387	9
Teacher emotional literacy motivation score	11.00	10.50	3.464	11	12.45	14.00	3.616	9
Teacher emotional literacy empathy score	13.36	14.50	2.790	8	13.73	16.00	2.936	7
Teacher emotional literacy social skills score	14.43	14.50	1.651	4	15.18	16.00	1.779	6
Teacher emotional literacy total score	63.29	63.50	11.125	29	68.18	77.00	13.280	36
Pupil BG Steem score	17.64	18.00	1.692	6	17.91	18.00	1.578	6
Parent SDQ emotional difficulties score	2.07	2.00	2.129	6	3.36	3.00	2.501	8
Parent SDQ behaviour difficulties score	1.07	.50	1.492	4	1.82	1.00	2.089	7
Parent SDQ hyperactivity score	4.29	3.00	2.946	9	3.82	3.00	2.822	10
Parent SDQ peer problems score	1.07	.00	1.542	4	2.27	2.00	2.195	8
Parent SDQ pro-social behaviour score	8.64	9.00	1.598	6	7.82	9.00	2.601	7
Parent SDQ total difficulties score	8.43	5.50	7.024	23	11.27	9.00	7.240	24
Teacher SDQ emotional difficulties score	2.29	2.00	2.091	7	1.18	.00	2.639	9
Teacher SDQ behaviour difficulties score	1.14	1.00	1.351	4	.91	.00	2.386	8
Teacher SDQ hyperactivity score	2.93	2.00	3.100	9	2.36	.00	2.767	6
Teacher SDQ peer problems score	1.50	.50	1.871	5	.82	.00	1.779	6
Teacher SDQ pro-social behaviour score	7.79	8.50	1.762	5	8.45	9.00	1.809	6
Teacher SDQ total difficulties score	7.86	6.50	6.087	17	5.27	1.00	7.101	20

Table 4.6 The mean, median, standard deviation and range of scores for the New Beginnings evaluation at follow-up

Dependent variable	Intervention group				Control group			
	M	Md	SD	R	M	Md	SD	R
Pupil emotional literacy total score	75.50	74	8.576	30	78.91	76	12.145	39
Parent emotional literacy self-awareness score	12.11	12	1.900	6	13.00	12	1.732	3
Parent emotional literacy self-regulation score	12.11	12	4.256	14	16.00	18	3.464	6
Parent emotional literacy motivation score	12.44	12	3.909	14	15.33	14	3.215	6
Parent emotional literacy empathy score	13.89	16	3.919	11	19.33	19	.577	1
Parent emotional literacy social skills score	17.89	17	2.028	6	18.00	18	2.000	4
Parent emotional literacy total score	68.44	71	14.196	43	81.67	80	9.609	19
Teacher emotional literacy self-awareness score	11.92	13	3.204	8	12.00	13	2.720	9
Teacher emotional literacy self-regulation score	9.75	7	5.446	12	12.27	14	4.519	12
Teacher emotional literacy motivation score	12.00	12	2.594	8	11.64	11	3.042	9
Teacher emotional literacy empathy score	10.58	10	3.801	10	14.09	16	3.360	9
Teacher emotional literacy social skills score	13.83	14	2.038	5	13.45	12	2.162	5
Teacher emotional literacy total score	56.92	49	16.681	42	63.00	60	12.182	37
Pupil BG Steem score	17.50	18	1.446	5	16.36	16	1.629	5
Parent SDQ emotional difficulties score	2.30	2	1.767	5	3.75	4	2.630	5
Parent SDQ behaviour difficulties score	3.10	3	2.644	8	.75	1	.500	1
Parent SDQ hyperactivity score	4.00	3	2.625	8	3.75	5	2.630	6
Parent SDQ peer problems score	2.40	2	2.459	7	1.50	2	1.291	3
Parent SDQ pro-social behaviour score	8.50	9	1.269	4	9.25	10	.957	2
Parent SDQ total difficulties score	11.80	12	7.285	19	9.75	12	5.315	12
Teacher SDQ emotional difficulties score	1.25	0	1.658	4	2.73	3	2.724	8
Teacher SDQ behaviour difficulties score	3.67	4.5	3.339	9	.91	0	1.640	5
Teacher SDQ hyperactivity score	4.92	6	4.033	10	2.73	2	2.573	8
Teacher SDQ peer problems score	2.33	1	3.200	10	1.18	1	1.168	3
Teacher SDQ pro-social behaviour score	6.33	5	2.640	7	8.27	10	2.149	5
Teacher SDQ total difficulties score	12.17	15	10.495	33	7.55	8	5.803	16

Table 4.7 The mean, median, standard deviation and range of scores for the Getting On and Falling Out evaluation at pre-test

Dependent variable	Intervention group				Control group			
	M	Md	SD	R	M	Md	SD	R
Pupil emotional literacy total score	77.42	74.5	9.030	27	79.36	77.0	13.411	43
Parent emotional literacy self-awareness score	12.67	14.0	2.309	4	12.83	13.0	.753	2
Parent emotional literacy self-regulation score	15.00	17.0	6.245	12	12.67	12.5	1.862	5
Parent emotional literacy motivation score	14.33	15.0	6.028	12	13.33	13.0	1.633	4
Parent emotional literacy empathy score	16.67	20.0	5.774	10	15.67	16.0	1.506	4
Parent emotional literacy social skills score	19.00	20.0	1.732	3	16.50	17.0	2.510	6
Parent emotional literacy total score	77.67	86.0	21.733	41	71.00	68.5	6.197	17
Teacher emotional literacy self-awareness score	11.83	11.5	3.538	10	12.33	11.0	3.082	7
Teacher emotional literacy self-regulation score	10.00	8.0	5.292	12	12.78	15.0	4.438	12
Teacher emotional literacy motivation score	11.58	11.0	3.147	10	11.67	11.0	3.202	9
Teacher emotional literacy empathy score	11.75	9.5	3.621	8	13.78	15.0	3.073	8
Teacher emotional literacy social skills score	13.92	14.5	2.151	6	13.78	13.0	2.167	5
Teacher emotional literacy total score	59.17	50.5	16.067	41	64.33	63.0	14.422	37
Pupil BG Steem score	17.08	18.0	2.429	8	16.18	17.0	2.960	10
Parent SDQ emotional difficulties score	1.33	1.0	.577	1	4.17	4.5	2.317	6
Parent SDQ behaviour difficulties score	1.33	.0	2.309	4	2.33	1.5	2.251	6
Parent SDQ hyperactivity score	5.00	5.0	5.000	10	4.50	5.0	1.378	4
Parent SDQ peer problems score	.33	.0	.577	1	2.50	2.0	2.258	6
Parent SDQ pro-social behaviour score	8.67	10.0	2.309	4	8.00	8.5	2.280	6
Parent SDQ total difficulties score	8.00	7.0	6.557	13	13.50	14.0	5.683	16
Teacher SDQ emotional difficulties score	2.00	2.0	1.907	5	1.50	.0	2.619	7
Teacher SDQ behaviour difficulties score	3.50	4.0	3.344	9	1.00	.0	1.927	5
Teacher SDQ hyperactivity score	3.92	3.5	3.753	10	2.63	1.0	3.420	9
Teacher SDQ peer problems score	1.75	1.0	2.417	7	1.00	.0	1.604	4
Teacher SDQ pro-social behaviour score	7.08	6.5	2.193	6	7.63	8.0	2.722	7
Teacher SDQ total difficulties score	11.17	11.0	9.889	25	6.13	5.0	6.813	18

Table 4.8 The mean, median, standard deviation and range of scores for the Getting On and Falling Out evaluation at post-test

4.3 Inferential statistics

4.3.1 Between-group comparisons

Similarity of the groups at pre-test

It was important to test whether the experimental and control groups were similar to begin with. Since participants were not randomly allocated to groups, there was a possibility that any differences at post-test could be attributed to pre-test differences. However, since the control group had been selected for a small group SEAL intervention later in the school year, it was hoped that the groups would score similarly at pre-test.

A Mann-Whitney *U* test was performed for each of the dependent variables to investigate the pre-test similarity between experimental and control groups. For the New Beginnings intervention, there were no significant differences between the experimental and control group on any dependent variables at pre-test. This indicates that the participants' scores were similar to begin with. It does not, however, indicate that the groups were equivalent, due to the non-randomised group selection process (Cook and Campbell, 1979).

For the Getting On and Falling Out intervention, there were significant differences between the experimental and control group on 4 out of the 26 dependent variables, and no significant differences between groups on the other 22 dependent variables at pre-test. The results of the Mann-Whitney *U* test for the variables with a significant difference between the intervention and control group at pre-test were as follows. There was a significant

difference in the parent ELAI empathy score between the intervention group ($Md = 16, n = 9$) and control group ($Md = 19, n = 3$), $U = 1.000, z = -2.348, p = 0.019$. There was a significant difference in the teacher ELAI empathy score between the intervention group ($Md = 9.5, n = 12$) and control group ($Md = 16, n = 11$), $U = 30.000, z = -2.292, p = 0.022$. There was a significant difference in the teacher SDQ behaviour difficulties score between the intervention group ($Md = 4.5, n = 12$) and control group ($Md = 0, n = 11$), $U = 35.000, z = -2.026, p = 0.043$. There was a significant difference in the teacher SDQ pro-social behaviour score between the intervention group ($Md = 5, n = 12$) and control group ($Md = 10, n = 11$), $U = 32.000, z = -2.147, p = 0.032$.

This indicates that children in the intervention group were rated as lower on empathy by parents and teachers, higher on behaviour difficulties by teachers and lower on pro-social behaviour by teachers at pre-test. So for the Getting On and Falling Out intervention, the intervention and control groups were roughly similar on most measures at pre-test, but there were important differences in their levels of empathy and teacher-rated behaviour that affect the interpretation and analysis of the changes in these dependent variables.

Consideration was given to performing an analysis of covariance to take account of the differences between groups using pre-test scores, but since there were very few significant differences between groups at pre-test, and since the data were not normally distributed, this was not carried out.

Similarity of the groups at post-test

A Mann-Whitney *U* test was performed for each of the dependent variables to investigate the post-test similarity between experimental and control groups. For the New Beginnings intervention, there were no significant differences between the experimental and control group on any of the dependent variables at post-test. This indicates that the participants' scores were still similar at post-test. It does not, however, give any information about the change in scores from pre-test to post-test.

For the Getting On and Falling Out intervention, there were no significant differences between the experimental and control group on any of the dependent variables at post-test. This indicates that the participants' scores were similar at post-test. The reader will recall that at pre-test, children in the intervention group were rated as lower on empathy by parents and teachers, higher on behaviour difficulties by teachers and lower on pro-social behaviour by teachers. However, the Mann-Whitney *U* test revealed that these differences were not present at post-test. This indicates that either the children in the intervention group made an improvement in their scores on these measures, or the control group's scores had deteriorated.

In order to find out which explanation was correct, the median scores for these dependent variables for the intervention and control group at pre- and post-test, and were compared. The results are in table 4.9. They indicate that, for parent-rated empathy, the intervention group's scores improved

from pre to post-test whereas the control group's scores decreased, for teacher-rated empathy, the intervention group's scores stayed the same but the control group's scores declined, for teacher-rated behaviour difficulties, the intervention group had a small decrease in median score whereas the control group stayed the same from pre to post-test, and finally for teacher-rated pro-social behaviour, the intervention group's scores improved from pre to post-test but the control group's scores decreased. Again, these analyses do not give any information about whether the changes in scores from pre-test to post-test are statistically significant.

Name of dependent variable	Median score			
	Pre-test		Post-test	
	Intervention group	Control group	Intervention group	Control group
Parent ELAI - empathy	16	19	20	16
Teacher ELAI – empathy	9.5	16	9.5	15
Teacher SDQ – behaviour difficulties	4.5	0	4	0
Teacher SDQ – pro-social behaviour	5	10	6.5	8

Table 4.9: Median scores in the Getting On and Falling Out evaluation, for the dependent variables that were significantly different at pre-test but not at post-test

Summary of between-group comparisons

To summarise, for the New Beginnings intervention, there were no significant differences between the intervention and control group at either pre-test or

post-test on any of the dependent variables. For the Getting On and Falling Out intervention, children in the intervention group were rated as lower than the control group on empathy by parents and teachers, higher on behaviour difficulties by teachers and lower on pro-social behaviour by teachers at pre-test, but there were no significant differences between groups at post-test on any dependent variables. The line graphs in figures 4.2 to 4.5 illustrate how the intervention group ‘closed the gap’ with the control group at post-test.

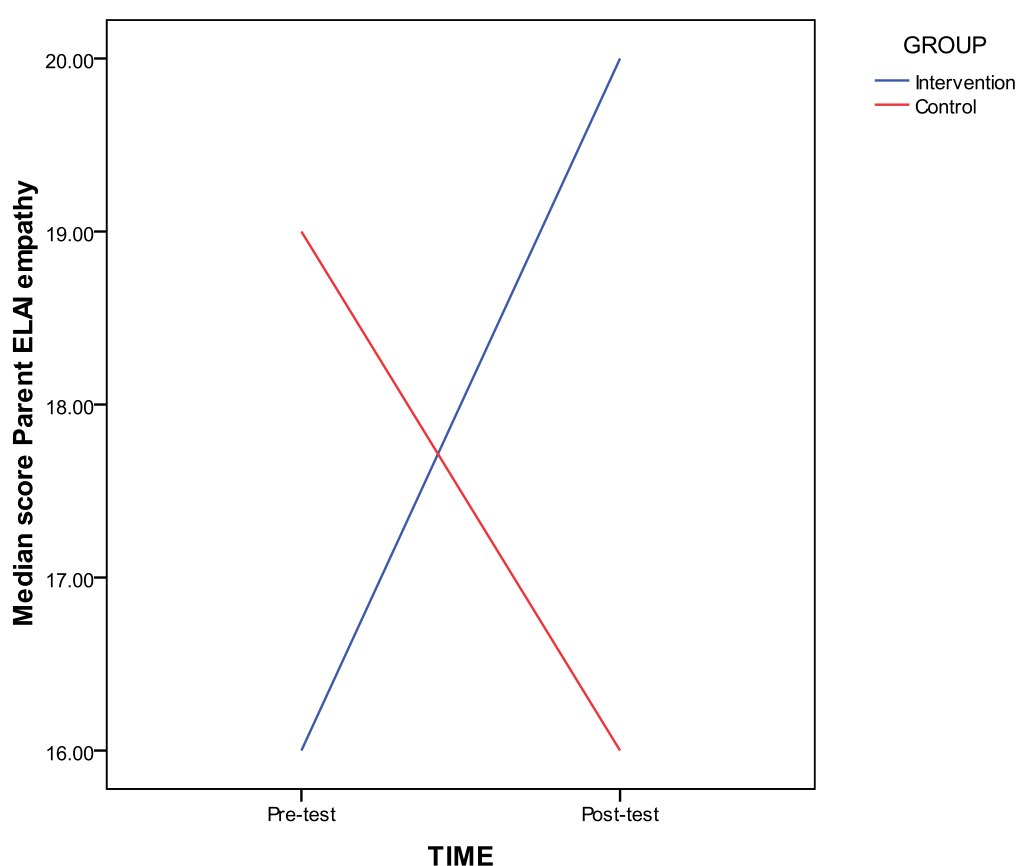


Figure 4.2 The differences between the intervention and control groups' scores at pre- and post-test on parent-rated empathy (Getting On and Falling Out)

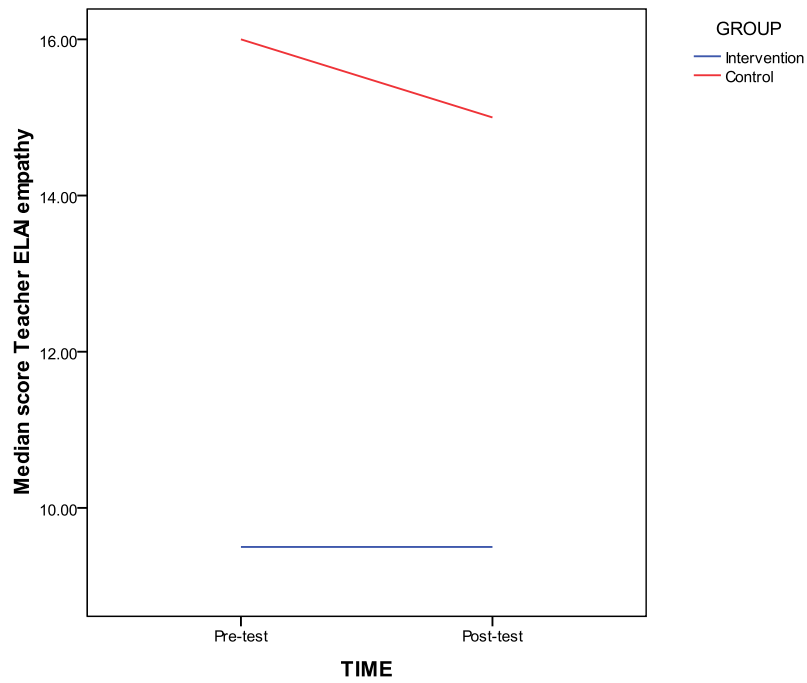


Figure 4.3 The differences between the intervention and control groups' scores at pre- and post-test on teacher-rated empathy (Getting On and Falling Out)

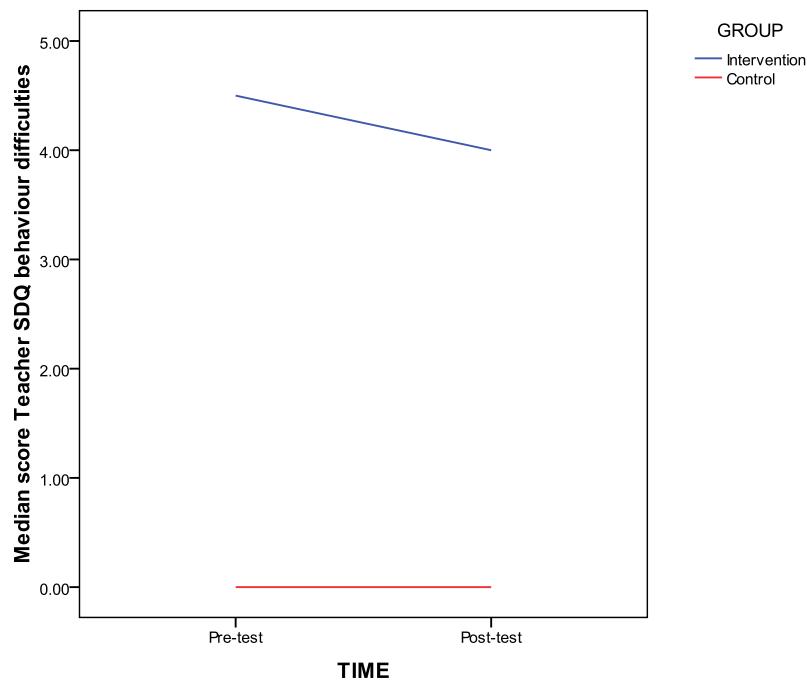


Figure 4.4 The differences between the intervention and control groups' scores at pre- and post-test on teacher-rated behaviour difficulties (Getting On and Falling Out)

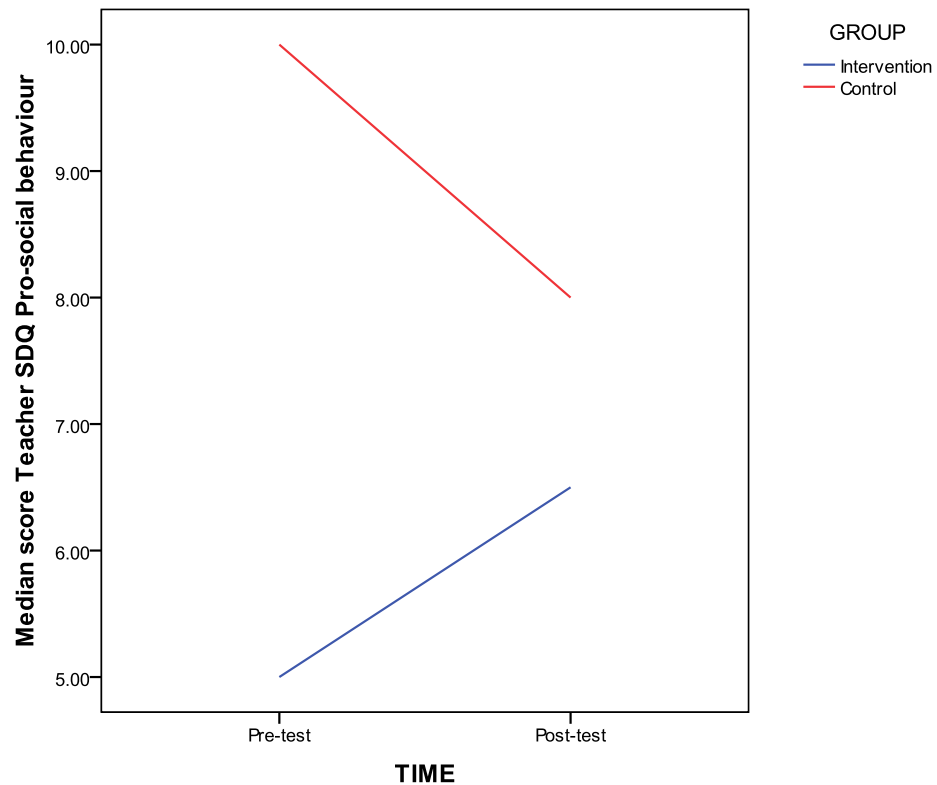


Figure 4.5 The differences between the intervention and control groups' scores at pre- and post-test on teacher-rated pro-social behaviour (Getting On and Falling Out)

4.3.2 Pre to post-test change in scores

New Beginnings

To compare pre to post-test differences in participants' scores, a series of Wilcoxon Signed Ranks Tests were performed. For the New Beginnings intervention group, the Wilcoxon Signed Ranks Tests revealed a significant difference between participants' scores at pre-test and post-test for pupil-rated total emotional literacy ($z = -2.672, p < 0.01$), teacher-rated self-awareness ($z = -2.201, p < 0.03$), teacher-rated total emotional literacy ($z = -2.414, p < 0.02$), teacher-rated hyperactivity ($z = -2.104, p < 0.04$) and teacher-rated total difficulties ($z = -2.110, p < 0.04$). There were no significant differences for the other dependent variables.

Closer examination of the data revealed that the changes were in socially desirable directions; increases in pupil-rated total emotional literacy, teacher-rated self-awareness and teacher-rated total emotional literacy and decreases in teacher-rated hyperactivity and teacher-rated total difficulties.

For the control group, the Wilcoxon Signed Ranks Tests revealed significant differences between participants' pre-test and post-test scores for pupil-rated total emotional literacy ($z = -2.767, p < 0.01$), teacher-rated self-awareness ($z = -3.007, p < 0.01$), teacher-rated motivation ($z = -2.389, p < 0.02$), teacher-rated empathy ($z = -2.979, p < 0.01$), teacher-rated social skills ($z = -2.793, p < 0.01$), teacher-rated total emotional literacy ($z = -3.365, p < 0.01$), teacher-rated emotional distress ($z = -2.687, p < 0.01$), teacher-rated pro-social behaviour ($z = -2.461, p < 0.02$), teacher-rated total difficulties ($z = -2.136, p < 0.04$). There were no significant differences on other dependent variables.

Closer examination of the data revealed that the changes were in socially desirable directions; increases in pupil-rated total emotional literacy, teacher-rated self-awareness, teacher-rated motivation, teacher-rated empathy, teacher-rated social skills, teacher-rated total emotional literacy and teacher-rated pro-social behaviour and decreases in teacher-rated emotional distress and teacher-rated total difficulties.

This information is summarised in table 4.10.

Dependent variable	Intervention group	Control group
	Result of Wilcoxon Signed Ranks Test	Result of Wilcoxon Signed Ranks Test
Pupil emotional literacy total score	$z = -2.672, p < 0.01$	$z = -2.767, p < 0.01$
Parent emotional literacy self-awareness score	Not significant	Not significant
Parent emotional literacy self-regulation score	Not significant	Not significant
Parent emotional literacy motivation score	Not significant	Not significant
Parent emotional literacy empathy score	Not significant	Not significant
Parent emotional literacy social skills score	Not significant	Not significant
Parent emotional literacy total score	Not significant	Not significant
Teacher emotional literacy self-awareness score	$z = -2.201, p < 0.03$	$z = -3.007, p < 0.01$
Teacher emotional literacy self-regulation score	Not significant	Not significant
Teacher emotional literacy motivation score	Not significant	$z = -2.389, p < 0.02$
Teacher emotional literacy empathy score	Not significant	$z = -2.979, p < 0.01$
Teacher emotional literacy social skills score	Not significant	$z = -2.793, p < 0.01$
Teacher emotional literacy total score	$z = -2.414, p < 0.02$	$z = -3.365, p < 0.01$
Pupil BG Steem score	Not significant	Not significant
Parent SDQ emotional difficulties score	Not significant	Not significant
Parent SDQ behaviour difficulties score	Not significant	Not significant
Parent SDQ hyperactivity score	Not significant	Not significant
Parent SDQ peer problems score	Not significant	Not significant
Parent SDQ pro-social behaviour score	Not significant	Not significant
Parent SDQ total difficulties score	Not significant	Not significant
Teacher SDQ emotional difficulties score	Not significant	$z = -2.687, p < 0.01$
Teacher SDQ behaviour difficulties score	Not significant	Not significant
Teacher SDQ hyperactivity score	$z = -2.104, p < 0.04$	Not significant
Teacher SDQ peer problems score	Not significant	Not significant
Teacher SDQ pro-social behaviour score	Not significant	$z = -2.461, p < 0.02$
Teacher SDQ total difficulties score	$z = -2.110, p < 0.04$	$z = -2.136, p < 0.04$

Table 4.10: The results of the Wilcoxon Signed Ranks Tests for pre- to post-test change (New Beginnings)

Getting On and Falling Out

For the Getting On and Falling Out intervention group, the Wilcoxon Signed Ranks Tests revealed a significant difference between participants' scores at pre-test and post-test for teacher-rated empathy ($z = -2.401, p < 0.02$), teacher-rated total emotional literacy ($z = -2.252, p < 0.03$) and teacher-rated pro-social behaviour ($z = -1.983, p < 0.05$). Closer examination of the data revealed that the changes were in socially desirable directions; increases in teacher-rated empathy, teacher-rated emotional literacy and teacher-rated pro-social behaviour. There were no significant differences for the other dependent variables. The same procedure was followed for the control group, and there were no significant differences when comparing pre- and post-test scores. This information is summarised in table 4.11.

Dependent variable	Intervention group	Control group
	Result of Wilcoxon Signed Ranks Test	Result of Wilcoxon Signed Ranks Test
Pupil emotional literacy total score	Not significant	Not significant
Parent emotional literacy self-awareness score	Not significant	Not significant
Parent emotional literacy self-regulation score	Not significant	Not significant
Parent emotional literacy motivation score	Not significant	Not significant
Parent emotional literacy empathy score	Not significant	Not significant
Parent emotional literacy social skills score	Not significant	Not significant
Parent emotional literacy total score	Not significant	Not significant
Teacher emotional literacy self-awareness score	Not significant	Not significant
Teacher emotional literacy self-regulation score	Not significant	Not significant
Teacher emotional literacy motivation score	Not significant	Not significant
Teacher emotional literacy empathy score	$z = -2.401, p < 0.02$	Not significant
Teacher emotional literacy social skills score	Not significant	Not significant
Teacher emotional literacy total score	$z = -2.252, p < 0.03$	Not significant
Pupil BG Steem score	Not significant	Not significant
Parent SDQ emotional difficulties score	Not significant	Not significant
Parent SDQ behaviour difficulties score	Not significant	Not significant
Parent SDQ hyperactivity score	Not significant	Not significant
Parent SDQ peer problems score	Not significant	Not significant
Parent SDQ pro-social behaviour score	Not significant	Not significant
Parent SDQ total difficulties score	Not significant	Not significant
Teacher SDQ emotional difficulties score	Not significant	Not significant
Teacher SDQ behaviour difficulties score	Not significant	Not significant
Teacher SDQ hyperactivity score	Not significant	Not significant
Teacher SDQ peer problems score	Not significant	Not significant
Teacher SDQ pro-social behaviour score	$z = -1.983, p < 0.05$	Not significant
Teacher SDQ total difficulties score	Not significant	Not significant

Table 4.11: The results of the Wilcoxon Signed Ranks Tests for pre- to post-test change (Getting On and Falling Out)

Summary of within-group comparisons

To summarise, for the New Beginnings intervention, there were significant improvements for the intervention group on five of the dependent variables from pre-test to post-test. However, this was complicated by the improvements on nine of the dependent variables that were seen in the control group over the same period.

For the Getting On and Falling Out intervention, results were more straightforward, with improvements on three dependent variables from pre-test to post-test for the intervention group, and no significant improvements for the control group.

For both interventions, there were no significant improvements on most of the dependent variables.

4.3.3 Analysis of follow-up measures

For New Beginnings, data were collected half a term after the children completed the intervention. Two questions can be addressed with this data. Firstly, did the children maintain post-test gains at follow-up? Secondly, were there any improvements at follow-up that were not present at post-test?

A series of Wilcoxon Signed Ranks Tests were performed on the five dependent variables that were significantly different from the pre-test at

post-test, in order to investigate whether the New Beginnings intervention group's gains at post-test had been maintained at follow-up. The results for pupil-rated total emotional literacy, teacher-rated self-awareness and teacher-rated total emotional literacy were not significantly different at follow-up compared with post-test, indicating that the gains had been maintained. However, for teacher-rated hyperactivity ($z = -2.309, p < 0.03$) and teacher-rated total difficulties ($z = -2.330, p < 0.03$), the scores were significantly worse at follow-up than at post-test. Further Wilcoxon Signed Ranks Tests showed that there were no significant differences for teacher-rated hyperactivity and teacher-rated total difficulties between pre-test and follow-up, indicating that the decline from post-test to follow-up resulted in scores that were similar to, rather than worse than, baseline levels. The same analysis was not completed for the gains that the control group made, since they received the Getting On and Falling Out intervention between post-test and follow-up.

Wilcoxon Signed Ranks Tests were then performed for the remaining variables to assess post-test to follow-up change. There was a significant difference for teacher-rated behaviour difficulties ($z = -2.041, p < 0.05$), however this change was in an undesirable direction, indicating that the teachers perceived more behaviour problems at follow-up than post-test. A further Wilcoxon Signed Ranks Test was performed, which showed that the scores at pre-test and follow-up were not significantly different for teacher-rated behaviour difficulties, indicating that the follow-up scores were similar to baseline levels.

The final set of comparisons was between the New Beginnings intervention group's pre-test and follow-up scores. Wilcoxon Signed Ranks Tests were performed for all dependent variables. There were significant differences between pre-test and follow-up scores for pupil-rated total emotional literacy ($z = -2.118, p < 0.04$), teacher-rated empathy ($z = -2.112, p < 0.04$) and teacher-rated total emotional literacy ($z = -2.206, p < 0.03$). The gains for pupil- and teacher-rated total emotional literacy had also been present at post-test and were not significantly higher than this level at follow-up. However, the gain in teacher-rated empathy was a new finding. There were no significant differences between pre and post-test and between post-test and follow-up, but the increase from pre-test to follow-up was significant. However, it was not possible to compare this finding with the control group, since any gains could have been attributed to the intervention they had received between the post-test and follow-up data collection points.

4.3.4 Analysis of sub-groups

Firstly, the data were split according to whether the participants were targeted as needing the intervention, or had been chosen to take part as role models, as the role models in Humphrey et al's (2008) did not show any significant effects of the intervention. Secondly, the data were split according to gender, as boys' and girls' social and emotional development may vary. Consideration was given to performing a mixed between-within subjects analysis of variance, but this idea was rejected since the assumption of normal

distribution was violated and splitting the participants into sub-groups would lead to small group sizes, particularly for Getting On and Falling Out.

New Beginnings

Wilcoxon Signed Ranks Tests were performed for targeted children in the intervention group. There were significant differences between pre- and post-test scores on pupil-rated total emotional literacy ($z = -2.190, p < 0.03$) and teacher-rated total emotional literacy ($z = -2.102, p < 0.04$). There were no significant differences for role model children in the intervention group.

For targeted children in the control group, there were significant differences between pre-test and post-test scores on pupil-rated total emotional literacy ($z = -2.513, p < 0.02$), teacher-rated self-awareness ($z = -2.698, p < 0.01$), teacher-rated empathy ($z = -2.979, p < 0.01$), teacher-rated social skills ($z = -2.550, p < 0.02$), teacher-rated total emotional literacy ($z = -3.009, p < 0.01$), teacher-rated emotional difficulties ($z = -2.39, p < 0.02$), teacher-rated pro-social behaviour ($z = -2.136, p < 0.04$) and teacher-rated total difficulties ($z = -2.094, p < 0.04$). There were no significant differences for role model children in the control group.

Wilcoxon Signed Ranks Tests were performed for the boys in the intervention group. There were significant differences between pre-test and post-test scores on pupil-rated total emotional literacy ($z = -2.697, p < 0.01$), teacher-rated total emotional literacy ($z = -2.007, p < 0.05$), parent-rated pro-social

behaviour ($z = -2.041, p < 0.05$) and teacher-rated hyperactivity ($z = -1.997, p < 0.05$). For girls in the intervention group there were no significant differences.

For boys in the control group, there were significant differences between pre-test and post-test scores on pupil-rated total emotional literacy ($z = -2.179, p < 0.03$), teacher-rated empathy ($z = -2.428, p < 0.02$) and teacher-rated total emotional literacy ($z = -2.316, p < 0.03$). For girls in the control group, there were significant differences between pre- and post-test scores on teacher-rated self-awareness ($z = -2.388, p < 0.02$), teacher-rated motivation ($z = -2.081, p < 0.04$), teacher-rated social skills ($z = -2.456, p < 0.02$), teacher-rated total emotional literacy ($z = -2.524, p < 0.02$) and teacher-rated emotional difficulties ($z = -2.041, p < 0.05$).

All statistically significant differences were in a socially desirable direction.

Getting On and Falling Out

Wilcoxon Signed Ranks Tests were performed for targeted children in the intervention group. The tests were not performed for the parent ELAI and parent SDQ measures due to insufficient numbers. For the other dependent variables, there was a significant difference between pre-test and post-test scores on teacher-rated empathy ($z = -2.232, p < 0.03$). For role model children in the intervention group, targeted children in the control group and role model children in the control group there were no significant differences.

Wilcoxon Signed Ranks Tests were performed for the boys in the intervention group. There was a significant difference between pre-test and post-test scores on teacher-rated empathy ($z = -2.041, p < 0.05$). For girls in the intervention group there was a significant difference between pre-test and post-test scores on teacher-rated overall difficulties ($z = -2.032, p < 0.05$). For girls and boys in the control group, there were no significant differences.

All significant differences were in a socially desirable direction.

4.4 Effect sizes

Finally, effect sizes were calculated where statistically significant differences were found between pre- and post-test scores for the intervention groups. Effect sizes were calculated by using the formula suggested by Pallant (2007) for the Wilcoxon Signed Ranks Test: dividing the z value by the square root of the number of cases. Then, to facilitate comparison with previous research, this was converted to Cohen's d using the conversion table provided by Dunst et al (2004). See table 4.12 for the results.

Intervention	Dependent variable	Effect size r	Effect size Cohen's d	Description
New Beginnings	Pupil ELAI – total emotional literacy	0.36	0.77	Medium-large
	Teacher ELAI – self-awareness	0.34	0.72	Medium-large
	Teacher ELAI – total emotional literacy	0.37	0.80	Large
	Teacher SDQ – hyperactivity	0.35	0.75	Medium-large
	Teacher SDQ – total difficulties	0.35	0.75	Medium-large
Getting On and Falling Out	Teacher ELAI – empathy	0.49	1.12	Large
	Teacher ELAI – total emotional literacy	0.46	1.04	Large
	Teacher SDQ – pro-social behaviour	0.40	0.87	Large

Table 4.12: Effect sizes for the significant changes for pre- to post-test scores for the intervention groups

The same calculations were then performed for the New Beginnings intervention group's pre-test to follow-up gain, the pre- to post-test gains made by targeted children for New Beginnings and Getting On and Falling Out intervention groups and for gains that were specific to boys and girls in the intervention groups for New Beginnings and Getting On and Falling Out. The results are in table 4.13.

It is important to highlight that effect sizes were calculated using the results of Wilcoxon Signed Ranks Tests, which are within-subject comparisons, therefore no comparison with the control group has been factored in.

Intervention	Dependent variable	Effect size r	Effect size Cohen's d	Description
New Beginnings	Teacher ELAI – empathy (pre-test to follow-up)	0.33	0.70	Medium-large
	Pupil ELAI – total emotional literacy (pre-test to post-test, targeted children only)	0.33	0.70	Medium-large
	Teacher ELAI – total emotional literacy (pre-test to post-test, targeted children only)	0.36	0.77	Medium-large
	Pupil ELAI – total emotional literacy (pre-test to post-test, boys only)	0.51	1.19	Large
	Teacher ELAI – total emotional literacy (pre-test to post-test, boys only)	0.40	0.87	Large
	Parent SDQ – pro-social behaviour (pre-test to post-test, boys only)	0.44	0.98	Large
	Teacher SDQ – hyperactivity (pre-test to post-test, boys only)	0.45	1.01	Large
Getting On and Falling Out	Teacher ELAI – empathy (pre-test to post-test, targeted children only)	0.56	1.12	Large
	Teacher ELAI – empathy (pre-test to post-test, boys only)	0.59	1.04	Large
	Teacher ELAI – total emotional literacy (pre-test to post-test, girls only)	0.59	0.87	Large

Table 4.13: Effect sizes for intervention groups' significant changes from pre-test to post-test for sub-groups, and for pre-test to follow-up gains

4.5 Intervention integrity

As described in section 3.5.3, the researcher developed a tool to assess the level of fidelity to the government guidance, which was completed for each school during a visit to observe a group-work session and interview the group leader. See Appendix F for a blank copy of the schedule and table 4.14 for the results of this investigation. The results are examined in the discussion.

Criterion	School						Total
	1	2	3	4	5	6	
Welcome/ check in activity	✓	✓	✓	✓	✓	✓	6/6 (100%)
Warm up activity	✓	✗	✓	✗	✓	✓	4/6 (67%)
Reminder of group aims/rules	✓	✓	✓	✗	✓	✓	5/6 (83%)
Reference to work from previous sessions or progress in the past week	N/A	✓	✓	✓	N/A	N/A	3/3 (100%)
Learning outcome shared	✓	✓	✗	✓	✓	✓	5/6 (83%)
Core activity	✓	✓	✓	✓	✓	✓	6/6 (100%)
Chance to review and reflect	✗	✗	✓	✗	✗	✗	1/6 (17%)
Task to carry out between sessions	✗	✗	N/A	✗	✓	✓	2/5 (40%)
Relaxation activity	✓	✓	✓	✗	✓	✓	5/6 (83%)
40 minute session or longer	✓	✓	✗	✓	✓	✓	5/6 (83%)
Access to supervision arrangements	✗	✓	✓	✓	✓	✓	5/6 (83%)
Weekly meeting with class teacher	✗	✗	✗	✗	✗	✗	0/6 (0%)
Group follows same SEAL theme as class	✓	✓	✓	✓	✓	✓	6/6 (100%)
At least six sessions	✓	✓	✓	✓	✗	✗	4/6 (67%)
Interruption-free room	✓	✓	✓	✓	✓	✓	6/6 (100%)
Children's progress shared with parents/carers	✓	✗	✓	✗	✓	✓	4/6 (67%)
Children told why they were chosen for the group	✗	✗	✗	✗	✗	✗	0/6 (0%)
Total	11/16 (69%)	11/17 (65%)	12/16 (75%)	9/17 (53%)	12/16 (75%)	12/16 (75%)	

Table 4.14: Fidelity to the intervention in the six schools

CHAPTER 5: DISCUSSION

5.1 Introduction to the discussion chapter

This chapter provides a fuller discussion of the findings of the study by summarising the results and interpreting these findings in the light of issues raised in the literature review and decisions made about research design. The thesis concludes with the professional implications of the study, areas for future research and personal reflections.

5.2 Discussion of results and data analysis

Since the results were very different for the two evaluations, this section is separated into the results for New Beginnings and the results for Getting On and Falling Out.

5.2.1 New Beginnings

The results of the New Beginnings evaluation are somewhat difficult to interpret. At pre-test, there were no statistically significant differences between the intervention and control group scores on any measures. However, this was also the case at post-test. An ideal result for a successful intervention would be to have similar pre-test scores, but for the intervention group to outperform the control group at post-test. This was not the case in the present study, which suggests that the New Beginnings intervention did not lead the intervention group to make gains over the control group.

The within-subject changes in score from pre- to post-test revealed statistically significant improvements for the intervention group on five variables: pupil-rated emotional literacy and teacher-rated self-awareness, total emotional literacy, hyperactivity and total difficulties. The effect sizes for these changes were medium to large. This would have been a good result for the intervention had the control group scores remained similar from pre- to post-test. However, the control group made statistically significant gains on four out of the five variables that the intervention group had improved on (the exception was teacher-rated hyperactivity) and also made statistically significant gains on five further variables: teacher-rated motivation, empathy, social skills, emotional difficulties and pro-social behaviour. Thus, it appears that the children in the control group made more of an improvement in teacher-rated emotional literacy than those who had the New Beginnings small group intervention. However, bearing in mind the post-test comparisons, this difference was not statistically significant.

The analysis of follow-up measures indicated that the gains made by the intervention group at post-test were maintained at follow-up for pupil-rated total emotional literacy, teacher-rated self-awareness and teacher-rated total emotional literacy. The post-test gains for teacher-rated hyperactivity and total difficulties were not maintained at follow-up and returned to levels similar to those found at pre-test. It was not possible to assess whether the control group's gains had been maintained as they received the Getting On

and Falling Out intervention between post-test and follow-up, so it would be unclear what any maintenance of gains was caused by. Another finding was that the intervention group improved on teacher-rated empathy from pre-test to follow-up and this had not been present between pre- and post-test. A possible explanation is that this was a delayed result of the New Beginnings intervention, although there are other possible explanations for this result.

When the data were analysed according to whether the participants were targeted as in need of intervention or as role models for the group, there were statistically significant improvements for the targeted pupils but not the role model pupils. When the data were analysed according to gender, in the intervention group, boys made gains on four variables whereas girls did not make any gains. However, in the control group boys made gains on three variables and girls made gains on five variables.

Overall, the results indicated that children in the control group made more gains in their social and emotional competence than children who had the New Beginnings intervention, but that the differences between groups at post-test were not statistically significant. This pattern of results is difficult to explain, but possible reasons are that the intervention was ineffective, that the control group's exposure to wave one SEAL caused their increased scores (this is unlikely since the intervention group were also exposed to SEAL at wave one) or that by being selected to go on a waiting list for a small group SEAL intervention the children in the control group were more 'visible' to

teachers and therefore their improvements were more noticeable (this is unlikely since the intervention group would also have been highlighted to teachers). Another possibility is interference from the simultaneous wave one SEAL intervention. It could be argued that New Beginnings occurs at a time (the start of the school year) when social and emotional issues are prioritised, so it might be that the wave two input does not add as much value to 'quality first teaching,' compared with other SEAL themes where wave one may not be given as much priority. This would add weight to the argument that social and emotional skills should be both 'caught' and 'taught' that was introduced in section 2.5. However, the data shed no light on this speculation and further research would be needed to test this hypothesis.

5.2.2 Getting On and Falling Out

The results of the Getting On and Falling Out evaluation are more straightforward, although not in the way explained above (similar scores at pre-test with the intervention group outperforming the control group at post-test). In this case, the intervention group had significantly worse scores than the control group at pre-test on parent-rated empathy and teacher-rated empathy, behaviour difficulties and pro-social behaviour. However, at post-test, there were no significant differences between the intervention and control group on these (or any other) tests. Possible interpretations of these results are that the intervention led the intervention group's scores to improve or stay the same whereas the control group's scores stayed the same or declined over the same period. These interpretations suggest that Getting

On and Falling Out had either an enhancing or protective effect. However this could be a simplistic explanation given the uncertainties in social research, and the reader is advised to consult section 5.4 for a range of competing explanations for the results, including the possibility of the intervention group's scores regressing to the mean.

The within-subject changes in score from pre- to post-test revealed statistically significant improvements for the intervention group on three variables: teacher-rated empathy, total emotional literacy and pro-social behaviour. Unlike the New Beginnings evaluation, there were no corresponding improvements for the control group. This could indicate that the Getting On and Falling Out intervention had a positive effect on participants, since the likelihood of the gains occurring by chance was very small. Also, the effect sizes for these improvements were large. However, it is not possible to make a definitive statement about causation, due to competing explanations for the results.

When the data were analysed according to whether the participants were targeted as in need of intervention or as role models for the group, there were statistically significant improvements for the targeted pupils but not the role model pupils. When the data were analysed according to gender, in the intervention group, boys and girls made a gain on one variable each, whereas boys and girls in the control group did not make any gains.

Overall, the results indicated that children who had the Getting On and Falling Out intervention made gains in their social and emotional competence (in particular in empathy and behaviour) in comparison to the control group who did not make any gains over the same period of time.

5.2.3 Issues of data analysis

The results discussed above are complicated by some issues with analysing the data, and these are now explained. Robson (2002) and Reichardt (1979) discuss analysing pre-test post-test non-equivalent groups designs, however the approaches that they suggest are all parametric tests whereas in this study the participants were not randomly allocated to conditions and some data were not normally distributed, therefore such tests were inappropriate. The non-parametric tests compared the intervention and control groups to each other at pre-test and post-test and compared within-group differences at pre- and post-test, but could not examine the interaction between time and group, as an analysis of variance could have done for example.

Another issue is effect size. The researcher calculated effect size in order not to rely solely on significance testing and to facilitate comparison with previous research. However, there are two reasons why these effect sizes should be treated with extreme caution. Firstly, effect size relies on group means; but the mean is not the best measure of central tendency in data that are not normally distributed, as was the case for much of the data in this study. Secondly, effect sizes for intervention groups are usually computed in relation

to a control group. However, in this case, effect sizes were calculated using results from Wilcoxon Signed Ranks Tests, which are within-subject tests.

Mertens and McLaughlin (2004) discuss dealing with the effects of a sample that is small and not randomly selected. They advise using non-parametric statistics, reporting effect sizes, replicating the study and discussing competing explanations and limitations of the study in the write-up. Non-parametric statistics and effect sizes have been discussed, replication was not feasible due to time constraints (New Beginnings and Getting On and Falling Out interventions would not be due to run in schools again until the following autumn) but in some ways the study is a replication of Humphrey et al's (2008) study. The final piece of advice is addressed in the following section.

5.3 Discussion of results in relation to the research questions

5.3.1 The effects on children's social and emotional outcomes

The research questions relevant to this aspect were:

What are the effects of the targeted (wave two) Primary SEAL small group intervention (silver set materials) on social and emotional outcomes for children in lower schools in Central Bedfordshire?

- Do pupils rate themselves as more skilled after the intervention?
- Do parents rate pupils as more skilled after the intervention?
- Do teachers rate pupils as more skilled after the intervention?

- Are the results more positive immediately after the intervention or several weeks later?

Pupil self-report

The hypothesis was:

Pupils rate themselves more favourably on social and emotional measures as a result of the small group SEAL intervention

This hypothesis was not supported. For the Getting On and Falling Out intervention, there was not enough evidence to reject the null hypothesis. Therefore, the increase in the intervention group's score from pre- to post-test on the pupil version of the ELAI might be explained by chance. For New Beginnings, there was an improvement in scores and enough evidence to reject the null hypothesis for the intervention group. However, the same result occurred for the control group, meaning that the intervention group's increases could be explained by the unknown factor that caused the control group's scores to rise, rather than the New Beginnings intervention.

Parent ratings

The hypothesis was:

Parents rate their children more favourably on social and emotional measures as a result of the small group SEAL intervention

This hypothesis was not supported. For the New Beginnings intervention, the intervention and control group had similar scores on parent versions of the ELAI and SDQ at pre- and post-test. Where increases in scores occurred from pre- to post-test for the intervention group, there was not enough evidence to reject the null hypothesis, meaning that the change could have occurred by chance. There was some evidence that scores improved for boys in the intervention group on parent-rated pro-social behaviour and that this did not occur for boys in the control group, however, caution should be exercised due to the small number of parent ratings for this variable (12 at pre-test and 10 at post-test). For Getting On and Falling Out, the intervention group scored significantly lower than the control group on parent-rated empathy at pre-test but not at post-test. However, there are several reasons why this cannot be taken as evidence to support the intervention. Firstly, the pre- to post-test change on this variable did not reach statistical significance, meaning that it could have occurred by chance. Secondly, the change could have occurred due to statistical regression. Thirdly, the change could have occurred due to an interaction between selection and the intervention. Finally, the number of parental responses was extremely low for this variable (9 for the intervention group and 3 for the control group).

Teacher ratings

The hypothesis was:

Teachers rate their pupils more favourably on social and emotional measures as a result of the small group SEAL intervention

This hypothesis was supported for some variables, but not for most. For New Beginnings, teachers rated pupils more favourably after the intervention but not in comparison with the control group. For Getting On and Falling Out, the intervention group had worse scores than the control group at pre-test on teacher-rated empathy, behaviour difficulties and pro-social behaviour but not at post-test. However this may have occurred due to statistical regression, and the change in behaviour difficulties was not significantly different from pre- to post-test so may have occurred by chance. The intervention group also made improvements on teacher-rated empathy, total emotional literacy and pro-social behaviour from pre- to post-test, whereas the control group did not make any improvements. These gains are undermined by threats to the validity of the quasi-experimental design.

The New Beginnings SEAL theme aims to develop empathy in children and the Getting On and Falling Out theme aims to improve children's self-regulation or management of feelings. Therefore, it is interesting that the gains for Getting On and Falling Out were in the area of empathy rather than self-regulation, and that no gains for empathy were seen in the New Beginnings evaluation

between pre- and post-test, although there was a statistically significant improvement in teacher-rated empathy between pre-test and follow-up. See the next section for more information.

Medium-term follow-up ratings

The hypothesis was:

The results are more favourable several weeks after a small group SEAL intervention than immediately afterwards

This hypothesis was supported for one variable, but not for the 25 others. Follow-up data were collected for the New Beginnings evaluation. Scores for teacher-rated empathy were significantly different from pre-test to follow-up but not from pre-test to post-test. However, there was no control group with which to compare this pattern of results, as the control group had the Getting On and Falling Out intervention between post-test and follow-up. From post-test to follow-up, there were no significant improvements, some gains were maintained and some scores returned to baseline levels.

5.3.2 The effects of the intervention on children's self-esteem

The research question was:

Does the targeted (wave two) Primary SEAL small group intervention (silver set materials) have an effect on children's self-esteem?

The hypothesis was:

Pupils rate themselves more favourably on self-esteem measures as a result of the small group SEAL intervention

This hypothesis was not supported. For New Beginnings and Getting On and Falling Out, the intervention and control groups had similar scores to each other at both pre- and post-test, and there were no significant differences within the groups from pre- to post-test nor were there any significant differences for any of the sub-group analyses. Closer inspection of the results indicates that children's scores for self-esteem increased for New Beginnings and decreased for Getting On and Falling Out, but the differences were not statistically significant. This means that there is not enough evidence to reject the null hypothesis and therefore the results could be explained by chance. These findings were surprising, as it was hypothesised that children may make gains in areas broader than emotional literacy. It is particularly surprising for the role model participants, who were not perceived to need to make any improvements in emotional literacy and psychological adjustment but who may have improved their self-image by taking part in the groups.

5.3.3 The fidelity to government guidance on running the intervention

The research question was:

What is the level of fidelity to the government guidance on the targeted (wave two) Primary SEAL small group intervention (silver set materials)?

No hypothesis was stated as there was no previous research in this area. This question was addressed in a very different way to the other research questions. In order to find out more about the implementation of the SEAL small group interventions, the researcher devised a checklist using the government guidance on implementing SEAL as a basis. The checklist was then completed by visiting each school to observe a small group SEAL session and interview the group leader(s). The results were presented in table 4.14 in the results chapter. The results indicated that the level of adherence to the guidance by school staff varied from around half to three quarters of the factors identified. The most common factors adhered to were having a welcome activity, a core activity, a distraction-free room and following the same SEAL theme as the class and school (all 6 schools did this in their SEAL small group interventions). The most common factors not adhered to were a weekly meeting with the class teacher and telling the children why they had been chosen to be in the group (these did not happen in any of the 6 schools).

However, there are some difficulties with interpreting these results. Firstly, the SEAL materials are different from other emotional literacy interventions

that are more scripted because, as with other government-issued curriculum materials, they are designed to be used flexibly and adapted to the needs of the group. Therefore, a lower level of fidelity may not necessarily lead to a poorer quality intervention, and indeed may result in an intervention which is better suited to the unique needs of the group. Informal data that was not captured by the checklist included the observation by the group leader in school 2 that some of the materials were felt to be slightly advanced for their Year 3 pupils; therefore they supplemented some of the tasks with some activities from a book of social skills games for younger children. However, on the checklist, this school showed a fairly high level of fidelity to the guidance since they still followed the suggested format of the sessions.

Another difficulty with interpreting the results from the checklist is validity. The researcher designed the checklist to be as objective as possible, with questions that resulted in a clear yes or no answer (see Appendix F). However, for some of the factors the questions were not applicable, for example if the session observed was the first or last in an intervention. Also, there is a need for further information to be gathered on the reliability of the instrument, for example inter-rater reliability was not assessed.

5.4 Discussion of methodological implications and limitations

5.4.1 Threats to internal validity

As discussed in the methodology chapter, the chosen research design was the most realistic in an applied research context but was a compromise as it is

more vulnerable to threats to internal validity than designs such as RCTs. These threats to validity weaken the confidence in the experimental design, meaning that any effects at post-test could be attributed to factors other than the intervention being studied. In the present study, it would be tempting to conclude that the Getting On and Falling Out intervention caused the pre- to post-test gains in the experimental group but not the control group. However, Campbell and Stanley (1963) argue that it is essential for designers of quasi-experiments to be aware of alternative explanations for their results. The following section explores some competing explanations.

Campbell and Stanley (1963) argue that the non-equivalent control group design controls for the following threats to internal validity: history, maturation, testing, instrumentation, selection and mortality (see table 3.1). However, they suggest that selection-maturation interactions are a definite weakness and that regression is a possible source of concern in this design.

Regarding selection-maturation interactions, in the present study the intervention group were selected for the Getting On and Falling Out intervention and the control group were on the waiting list for the Going for Goals intervention. However, the needs of the intervention group might have been greater than those of the control group. For example, children might have been chosen for Getting On and Falling Out, not because of the specific content of the programme, but because they were in more urgent need of intervention. If this was the case (which is plausible given the pre-test

differences between the intervention and control groups on empathy and behaviour) the intervention group gains could have occurred as a result of the 'spontaneous remission' of a group with extreme scores, rather than due to receiving the intervention.

Regarding regression, Campbell and Stanley (1963) advise against attempting to control for pre-test differences between groups by matching if this is not accompanied by random assignment to conditions. This pitfall was avoided in the present study, however, regression to the mean is still a possible explanation for the Getting On and Falling Out findings. Barnett et al (2005) warn that regression to the mean is a problem in many repeated-measures studies as an extreme score is likely to be followed by a score that is closer to the mean. They recommend using an analysis of co-variance to overcome the effects of regression to the mean, but this could not be used with the data collected, therefore regression to the mean remains a plausible explanation for the increase in the intervention group's scores from pre- to post-test.

5.4.2 Threats to external validity

Campbell and Stanley (1963) suggest that testing-intervention interactions are a definite threat to external validity for non-equivalent control group designs, and that selection-intervention interactions and reactive arrangements are a possible source of concern in this design.

The testing-intervention interaction refers to the effect that the pre-test has on the effectiveness of the intervention, meaning that the benefits of the intervention cannot be generalised to participants who have not had the pre-test. However, Campbell and Stanley (1963) suggest that this is a particular threat in studies of attitude change and less of a threat in education, where assessment is more typical. This threat to external validity could be overcome by replicating the study with different outcome measures, as both this study and Humphrey et al's (2008) used the ELAI and SDQ.

The selection-intervention interaction is the likelihood that the selection of participants affects the results, meaning that the benefits of the intervention cannot be generalised to other participants or settings. Campbell and Stanley (1963) suggest that schools which agree to take part in research are not representative and are more likely to have higher staff morale, lower fear of inspection and more zeal than most schools. They advocate researchers being clear about how many schools were approached, as the author does in figure 3.1. Mertens and McLaughlin (2004) argue that, where random sampling from the target population has not occurred, inferences beyond the sample are affected. This is certainly relevant to the present study. A convenience sample was used for practical and ethical reasons, but this results in a lack of generalisability as participants may not reflect the wider population.

Reactive arrangements are the effects of the artificiality of the experiment, meaning that results cannot be generalised beyond the experimental

situation. Campbell and Stanley (1963) advise that these effects are worse in situations where the intervention or staff are unusual, but in this case both the intervention and the group leaders were part of normal school practice.

Another threat to generalisability in this study is sample size. Although there were some positive results for the Getting On and Falling Out intervention, this was based on an intervention group of 12 and control group of 11 pupils.

Although external validity is a concern, there are some reassurances. Firstly, replication is one way to improve generalisability and this study attempts to replicate Humphrey et al's (2008) evaluation. Secondly, since the study cannot be generalised beyond the sample, the participants have been described in detail so that readers can decide how similar they are to other populations or so that they can choose to replicate the study with a very different group of people.

5.4.3 Limitations of the study

A common limitation of educational research studies is their small scale. This is also the case in the present study, particularly for the Getting On and Falling Out evaluation and for sub-group analyses. However, hopefully one of the effects of the Development and Research project will be to aggregate trainee EPs' research, which might help to overcome this problem.

Another limitation is the lack of randomisation in sampling and allocating to groups. These have resulted in threats to the internal and external validity of the quasi-experiment, which affect its interpretation.

A further criticism could be that self-report and informant-report measures were used, rather than direct measures of behaviour (for instance role plays of social situations or direct observations of playground behaviour), as some researchers argue that direct measures are more objective. However direct measures concern only observable behaviour rather than cognitions and the problems of measuring social and emotional skills were discussed in section 3.4.2. Self-report measures were chosen over direct measures because they were felt to be more time-efficient and related to the theory of trait, rather than ability, EI. Also, Humphrey et al (2008) used role play measures in case study schools and found no effects.

Another issue with measurement is that self-awareness is a component of emotional literacy. Therefore, if a child improved their emotional literacy by becoming more self-aware, their score on that component may have increased but their overall emotional literacy might have decreased as a result of more realistic ratings in other areas. Since the pupil ELAI only gives total emotional literacy results, this may mask improvements in self-awareness.

A final criticism of the study is that it compared the SEAL small group interventions with a waiting list control group, rather than with an alternative

intervention. This decision was also made on practical grounds, as schools were unlikely to be able to staff two interventions simultaneously.

5.5 Discussion of results in relation to existing research findings

One of the main reasons for conducting the present research study was the lack of research into SEAL. However, this section attempts to integrate the findings with the little research that does exist.

Humphrey et al (2008) evaluated the New Beginnings intervention using a pre-test post-test control group design. For staff ratings of social and emotional outcomes (using the ELAI and SDQ) and for parent ratings on the ELAI, they found that both the intervention and control groups' scores improved over time, but there was no improvement relative to the control group. For parent SDQ ratings, no effects were detected for either group. For pupil self-report using the ELAI, they concluded that the intervention had a significant impact on scores compared with the control group, with an effect size of 0.44. The present study replicated this design, but failed to find a significant result for the intervention group over the control group for pupil-rated emotional literacy. In common with Humphrey et al (2008), the present study did not find any effects for parent and teacher ratings in comparison with the control group.

Humphrey et al (2008) evaluated the Getting On and Falling Out intervention using a single group phase change design. There was no control group, but

the authors argue that by taking two pre-test measures, the participants acted as their own control group using the baseline period. For teacher ELAI ratings, they found that children made statistically significant gains in empathy, motivation, self-awareness and social skills during the baseline phase and a statistically significant reduction in empathy in the intervention phase. For teacher SDQ ratings, scores decreased in both the baseline and intervention phases, although the decrease in the baseline phase was greater. For pupil self-report using the ELAI, scores for self-awareness, motivation, empathy and social skills were stable whereas the scores for self-regulation decreased during the baseline phase. During the intervention phase, there was a statistically significant increase in social skills. No analysis was performed on the parent questionnaires due to a low response rate (16 parents returned questionnaires at all three time points, out of 46 children). The present study did not replicate this design, for several reasons.

Firstly, Humphrey et al's (2008) baseline phase only had two measures, whereas it is preferable to identify a trend using at least three data collection points in an interrupted time-series design. Secondly, the logic of the design relies on a stable baseline period, which was not present for most ratings in the Humphrey et al study. The control group data from the New Beginnings evaluation could have been used to analyse Getting On and Falling Out on a single-group phase change design basis, however, the reader will recall from the results of the New Beginnings evaluation that the control group made gains from pre- to post-test, therefore there would not have been a stable

baseline phase had the design been analysed in this study. Finally, it was felt that a pre-test, post-test control group design was stronger for drawing conclusions. Unlike the Humphrey et al study which found very little evidence to support the Getting On and Falling Out intervention, the present study found gains for the experimental group relative to the control group particularly on teacher-rated empathy and pro-social behaviour.

The present study found quite different results from the Humphrey et al (2008) study, the only other evaluation of SEAL silver set interventions. Such contradictory results lead to uncertainty about the interventions' effects.

Downey and Williams (2009) found gains for all aspects of teacher-rated emotional literacy and parent-rated motivation, using the ELAI, for targeted children who participated in the Family SEAL intervention. However, there was no control group so it is possible that, like the New Beginnings evaluation in this study, these gains may also have been seen in a comparison group.

A similarity between the findings of this study and those of Humphrey et al (2008) and Downey and Williams (2009) is that parents perceived less impact of wave two SEAL interventions on children than teachers did. The present study found that children in the Getting On and Falling Out intervention group were rated lower than the control group on empathy at pre-test by parents but as similar at post-test, however the pre- to post-test change was not statistically significant. The rest of the positive effects found were for teacher

ratings. This is also similar to the results of Hallam et al's (2006) study of small group emotional literacy work in schools, which found significant results for teacher- but not parent-ratings using the SDQ. However, Parton and Manby (2009) found the opposite pattern, with parents' ratings indicating more of an improvement than teachers'.

A consistent finding of this study was the lack of impact on the role model pupils for the New Beginnings and Getting On and Falling Out interventions. Humphrey et al (2008) also found no effects on role models for the four interventions they evaluated. Downey and Williams (2009) found more effects for targeted pupils in the Family SEAL intervention, although there was a positive effect for teacher-rated self-awareness for the other pupils.

Another commonality with previous studies is the equivocal nature of the results. Overall, there were positive results for Getting On and Falling Out but for New Beginnings, the control group did slightly (although not statistically significantly) better than the intervention group. Humphrey et al (2008) also found some negative results alongside their positive findings. For example, there was a *reduction* in staff-judged empathy for Getting On and Falling Out. Similarly, Hallam et al (2006) found no significant change for pupils in Key Stage 1 in their evaluation of the wave one SEAL curriculum materials and found a small negative change for social skills and relationships in Key Stage 1 and increases in social skills and perceiving own emotions for Key Stage 2 for the pupil questionnaires in their evaluation of small group work.

A common theme in the existing research is a poor response rate from parents, which was also found in this study. In the present study, the response rates for parents were 73%, 54% and 55% at pre-test, post-test and follow-up for New Beginnings (or 90%, 67% and 75% excluding school 4) and 57% and 39% at pre-test and post-test for Getting On and Falling Out. Similarly, Humphrey et al (2008) started with a sample of 624 pupils, but only 295 parent questionnaires were returned at pre-test, 192 at post-test and 138 at follow-up, which gives response rates of 47%, 31% and 22% (Humphrey et al, 2008, do not report these response rates but rather express them as loss-to-follow-up rates, which are lower but obscure low response rates at pre-test). Hallam et al (2006) received 26 questionnaires from parents in their study, but do not comment on how many were sent out. However, given that questionnaire data were available for 9944 children at pre-test, the parental involvement in the evaluation seems very small indeed. Maddern et al (2004) also had difficulty with involving parents but found that they responded well to home visits from an assistant psychologist.

5.6 Strengths of the study

Consideration has been given to the limitations of the research, but there are some distinctive features and strengths which will now be outlined.

Firstly, this study has added to the under-researched area of the effectiveness of psychological interventions that are designed to promote social and

emotional outcomes such as emotional literacy, in UK schools. In particular it has added to the evidence on SEAL, which is a much neglected area of study given the amount of schools that are using it (63% of respondents to the questionnaire in this study, and 80% nationally according to Humphrey et al, 2008) and the amount of research produced (5 studies).

Secondly, this study used a control group design to answer the research questions on effectiveness. Commentators (for example Weare, 2004) have noted that there is a huge need for such controlled studies, as much of the research has been of pre-experimental quantitative design or used qualitative methods, which limits the internal and external validity. To the researcher's knowledge, this is the first pre-test post-test non-equivalent groups quasi-experiment on the Getting On and Falling Out intervention.

Another strength of the study is that a variety of respondents were involved: pupils, teachers and parents. In similar pieces of research, pupils' (for example Downey and Williams, 2009) and parents' (for example Hallam et al, 2006) views have been under-represented. However, such triangulation of data gives richer information that could be used to advance the field in future. For example, it would be useful to find out more about why parents perceive less of an impact of emotional literacy interventions in schools than teachers.

A further strength of the study is that the New Beginnings evaluation is a replication of research done by Humphrey et al (2008) into this intervention.

Since both this study and the Humphrey et al (2008) study are subject to threats to internal and external validity, it is important to try to overcome these threats through replication and thereby increase confidence in the results and the generalisability of the findings.

Although this study is a replication of the only other study on small group SEAL interventions (Humphrey et al, 2008), it can be distinguished from that study in several ways. Firstly, and as already mentioned, it used a control group design to assess the effectiveness of both interventions, rather than using a single-group phase change design for Getting On and Falling Out. Secondly, this study assessed fidelity to the government guidance on small group SEAL work by observing sessions and interviewing group leaders using a checklist. This was important because Humphrey et al's (2008) study and a previous study of SEAL in Bedfordshire (Bedfordshire County Council, 2008) raised concerns about how SEAL is being implemented. However, it would have been strengthened further had the inter-rater reliability of the checklist been assessed. Finally, this study has asked an original question about the effects of the small group SEAL interventions on children's self-esteem. The results of the study do not indicate that targeted social and emotional group interventions such as the SEAL silver set group work have any effect on children's self-esteem, as measured by the BG Steem, whereas the previous study (Humphrey et al, 2008) focused on measures of emotional literacy and psychological adjustment.

Further strengths of the study in terms of implications for EPs and the questions it raises are addressed in the next two sections.

5.7 Professional implications of the research

This section discusses the implications of the findings for applied EPs. The most obvious question is whether the small group SEAL work is an intervention that could be recommended to school staff by EPs seeking to promote evidence-based practice. The current findings, when viewed alone and when considered alongside the findings of Humphrey et al (2008), indicate that EPs should exercise caution in recommending this intervention as further research is needed. There was tentative support for Getting On and Falling Out for teacher ratings but a lack of evidence for New Beginnings.

Aside from evidence-based practice, another issue in recommending small group SEAL interventions is whether it is ethical to do so. At one stage, the author was concerned that the intervention group had been harmed by receiving the New Beginnings intervention as the control group had better outcomes. However, the lack of significant differences between the two groups at both pre and post-test would suggest otherwise.

Another set of results concerned the fidelity to government guidance. In this study, it ranged from 53% to 75%. This highlights a possible role for EPs, with their knowledge of the social psychology of running group work and the

research evidence on social and emotional interventions, to deliver, co-deliver or facilitate (on a consultative basis) the delivery of such interventions.

One possible interpretation of the lack of impact of the wave two New Beginnings intervention was that the wave one New Beginnings is effective and given a high priority at that stage in the academic year, and the wave two follow-up work may not add much value. If this is the case, it could undermine the use of wave two interventions and indicate that EPs' efforts should be focused on strengthening wave one provision, which would benefit all children, not just those who have been targeted for intervention.

Other implications of this research concern EPs or trainee EPs working as research practitioners. This study is a successful example of a trainee EP evaluating an intervention, which shows that this role can occur alongside more traditional EP activities.

The final implication is that more research into small group SEAL interventions is clearly needed, and EPs could be well placed to carry out this research, given their research skills, access to local authority personnel with responsibility for SEAL and existing relationships with schools. This research may take the form of replications of this type of group comparison study, gathering evidence on the intervention through single-subject research as part of routine casework practice or qualitative studies exploring the factors that facilitate or constrain the development of successful small group SEAL

interventions. The next section considers which aspects of small group SEAL interventions could be addressed by future research.

5.8 Future research

This study has raised several questions and also left unanswered questions. The following section describes these.

The research questions about the effectiveness of the intervention were not answered conclusively. This study was a part replication of Humphrey et al's (2008) investigation, the only other research into the SEAL silver set materials. Some of the results from this study contradicted the results of the previous study. This indicates that further research which replicates these studies is needed before the interventions can be supported or rejected with certainty.

Apart from replications, further research could be carried out that has an even stronger design for answering questions about the intervention's effectiveness. For example future research could include RCTs or could compare small group SEAL interventions to other social and emotional interventions, rather than using a waiting list control group.

A limitation of the present study's scope is that it only investigated New Beginnings and Getting On and Falling Out. Further research is needed on the other SEAL themes that have materials for wave two small group intervention.

Also, this study did not collect follow-up data for the children that had Getting On and Falling Out so further studies that do this would be of use.

More research is needed on the effects of the intervention on role model pupils. No effects were found for role model pupils in this study or in Humphrey et al's (2008) study. There is a clear theoretical basis from social cognitive theory for including role models in the groups, but so far no evidence to suggest a positive impact on these pupils, even on more general measures such as self-esteem. This raises ethical issues, for example is it ethical for them to spend time doing an intervention that gives them no clear benefit, when they could be benefitting from class teaching in the lessons they are missing? If further research is unable to provide ethical justification for including role model pupils, it may be necessary for the adult group leaders to become more explicit models of positive behaviours. However, the absence of role models from the group may adversely affect group composition and prevent targeted pupils from making progress, so this issue is a sensitive and controversial one.

The present study did not find any evidence to support the hypothesis that the small group SEAL interventions have an impact on more general measures of psychological wellbeing such as self-esteem, even for the role model pupils. Future research could investigate this hypothesis further by investigating other measures of self-esteem or other measures of psychological wellbeing such as self-image, confidence or resilience.

Further research into the effect of the intervention on parents' views would be beneficial. In this study, there was no impact on parents' views and further research may wish to replicate this aspect to see whether the same result is found, to answer research questions about whether there are any other effects on parents that were not measured by this study or to investigate the reasons for the lack of perceived parental impact.

5.9 Personal learning points

The author learned a lot from conducting the study, including some valuable lessons about conducting research that can be applied in future. These concerned methodology and real world practicalities.

An RCT, although less practical, would have been a better design for drawing firm conclusions from the results. In future, it might be better to run an RCT to answer similar research questions, although the feasibility of this design in applied settings would probably result in a reduced sample size.

A qualitative or constructivist approach was rejected due to the nature of the research questions. However, in future it might be useful to consider collecting qualitative data alongside quantitative data. Miller and Todd (2002) write that the researcher should not feel pushed towards a particular methodology and epistemology in evidence-based practice, but that there can be negotiation about methods and the nature of evidence. Miller and Todd

(2002) advocate a mixed methods approach to evaluation research so that the process and content of interventions can be considered alongside the outcomes. In the present study, this would have permitted the researcher to find out more about why parents perceived little impact. Another learning point was people's need not to be reduced to numbers. For example when the completed questionnaires from School 1 were returned, the group leader also provided quotes from the children on what they thought they had achieved or learned. The overall learning point was the importance of being clear about the questions that the researcher wants to answer.

Another factor that the researcher considered was the relevance of the research design to future research as an EP working with schools. It could be argued that EPs typically engage in single-case or action research in the school setting, but the author has learned that other research designs are feasible and the primary consideration should be the research questions.

The researcher learned a great deal about the practicalities of conducting research in the applied setting. There were unexpected events to adapt to, for example questionnaires not being returned on time, a delay in employing the family worker which led to the intervention starting later than planned in schools 5 and 6, changes in group composition in schools 2 and 4 and data being lost in school 4. The overall learning point was that matters that are very important to the researcher are often not of the utmost importance to others in the applied setting.

5.10 Conclusion

This thesis has reported on an investigation into the impact of primary SEAL small group work (silver set materials) on social and emotional outcomes for pupils. The author aimed to place the research in context in chapters one and two, explain the data collection and analysis in chapters three and four and finally discuss the results and implications in this final chapter. The results indicated that the New Beginnings intervention was not associated with any gains relative to a control group and the Getting On and Falling Out intervention was associated with improvements in teacher-rated empathy, total emotional literacy and pro-social behaviour. This study has added to the small evidence-base on the SEAL programme, but further research is needed.

REFERENCES

Ashkanasy, N. and Daus, C. (2005) Rumors of the death of emotional intelligence in organizational behavior are vastly exaggerated. *Journal of Organizational Behavior*, 26, 441-452.

Bandura, A. (1986) *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs: Prentice-Hall.

Barnett, A., van der Pols, J. and Dobson, A. (2005) Regression to the mean: what is it and how to deal with it. *International Journal of Epidemiology*, 34, 215-220.

Bedfordshire County Council (2008) *Monitoring Primary Social Emotional Aspects of Learning in Bedfordshire*. Unpublished report.

Borgers, N., Hox, J. and Sikkels, D. (2004) Response Effects in Surveys on Children and Adolescents: The Effect of Number of Response Options, Negative Wording and Neutral Mid-Point. *Quality and Quantity*, 38, 17-33.

British Psychological Society (2008) *Ethical Principles for Conducting Research with Human Participants*. Retrieved online December 2008 from <http://www.bps.org.uk/the-society/code-of-conduct/ethical-principles-for-conducting-research-with-human-participants.cfm#principles>

CACI (2010) Acorn classification. Retrieved online September 2010 from <http://www.caci.co.uk/acorn-classification.aspx>

Campbell, D. and Stanley, J. (1963) *Experimental and Quasi-Experimental Designs for Research*. Chicago: Rand McNally.

Cobb, N. (1992) *Adolescence: Continuity, Change and Diversity*. London: Mayfield Publishing Company.

Cohen, J. (1994) The Earth is Round ($p < .05$). *American Psychologist*, 49, 12, 997-1003.

Cohen, L., Manion, L. and Morrison, K. (2000) *Research Methods in Education (5th Edition)*. London: Routledge.

Cook, T. and Campbell, D. (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. London: Houghton Mifflin.

Coolican, H. (2009) *Research Methods and Statistics in Psychology (5th Edition)*. London: Hodder Education.

Coopersmith, S. (1967) *The Antecedents of Self-esteem*. San Francisco: Freeman.

Coppock, V. (2007) It's Good to Talk! A Multidimensional Qualitative Study of the Effectiveness of Emotional Literacy Work in Schools. *Children and Society*, 21, 405-419.

Craig, C. (2007) *The Potential Dangers of a Systematic, Explicit Approach to Teaching Social and Emotional Skills (SEAL)*. Centre for Confidence and Wellbeing. Retrieved online July 2009 from www.centreforconfidence.co.uk/docs/EI-SEAL_September_2007.pdf

Curtis, C. and Norgate, R. (2007) An Evaluation of the Promoting Alternative Thinking Strategies Curriculum at Key Stage 1. *Educational Psychology in Practice*, 23, 1, 33-44.

Daily Mail (no date) Search results for “social and emotional aspects of learning.” Retrieved online May 2010 from <http://www.dailymail.co.uk/home/search.html?searchPhrase=%22social+and+emotional+aspects+of+learning%22>

Dancey, C. and Reidy, J. (2007) *Statistics Without Maths for Psychology* (4th Edition). Harlow: Pearson Education.

Department for Education and Skills (2002) *Targeting support: managing NNS/NLS intervention programmes. A guide for headteachers and senior managers*. Retrieved online July 2009 from <http://nationalstrategies.standards.dcsf.gov.uk/node/154631>

Department for Education and Skills (2003) *Every Child Matters*. London: The Stationery Office.

Department for Education and Skills (2005) *Excellence and Enjoyment: social and emotional aspects of learning (Guidance)*. Retrieved online July 2009 from <http://nationalstrategies.standards.dcsf.gov.uk/node/65859>

Department for Education and Skills (2006a) *Excellence and Enjoyment: Social and Emotional Aspects of Learning. Key Stage 2 Small Group Activities*. Retrieved online December 2008 from <http://nationalstrategies.standards.dcsf.gov.uk/node/89188>

Department for Education and Skills (2006b) *The Origins of Primary SEAL*. Retrieved online October 2009 from http://nationalstrategies.standards.dcsf.gov.uk/node/164450?uc=force_uj

Department for Education and Skills (2006c) *The Waves of Intervention Model*. Retrieved online October 2009 from <http://nationalstrategies.standards.dcsf.gov.uk/node/46539>

Department for Education and Skills (2007a) *Intensifying Support Programme (ISP) and Social and Emotional Aspects of Learning (SEAL) Birchwood Junior School – April 2006*. Retrieved online July 2009 from

http://nationalstrategies.standards.dcsf.gov.uk/node/88432?uc=force_uj

Department for Education and Skills (2007b) *Social and Emotional Aspects of Learning for Secondary Schools: Guidance Booklet*. Retrieved online July 2009 from

http://nationalstrategies.standards.dcsf.gov.uk/node/157981?uc=force_uj

Department of Health and Department for Children Schools and Families (2007) *Introduction to the National Healthy Schools Programme*. Retrieved online April 2010 from

<http://resources.healthyschools.gov.uk/v/2e4b94b5-693b-46a0-88e1-9cbc00f1ac09?c=8d58bfca-39d1-4d7d-927a-9cb501033fb9>

Department of Health/Department for Education and Skills (2004) *Promoting emotional health and wellbeing through the National Healthy School Standard*. Retrieved online July 2009 from

http://nationalstrategies.standards.dcsf.gov.uk/node/155520?uc=force_uj

Downey, C. and Williams, C. (2009) *Seeing Eye-to-eye: Evaluation of Family SEAL, a home-school learning partnership*. Conference presentation from the International Congress for School Effectiveness and Improvement 5th January 2009. Retrieved online July 2009 from

<http://www.icsei2009.org/presentations/DowneyWilliams-FamSEAL.pdf>

Dunst, C., Hamby, D. and Trivette, C. (2004) Guidelines for calculating effect sizes for practice-based research syntheses. *Centerscope: Evidence-based Approaches to Early Childhood Development*, 3, 1-10.

Ecclestone, K. (2007) Resisting images of the 'diminished self': the implications of emotional wellbeing and emotional engagement in education policy. *Journal of Education Policy*, 22, 4, 455-470.

Ecclestone, K. and Hayes, D. (2008) *The Dangerous Rise of Therapeutic Education*. London: Routledge.

Ecclestone, K. and Hayes, D. (2009) Changing the subject: the educational implications of developing emotional well-being. *Oxford Review of Education*, 35, 3, 371-389.

Elias, M., Zins, J. and Weissberg, R., Frey, K., Greenberg, M., Haynes, N., Kessler, R., Schwab-Stone, M. and Shriver, T. (1997) *Social and Emotional Learning: Guidelines for Educators*. Alexandria, Virginia: ASCD.

Faupel, A. (Ed.)/Southampton City Council (2003) *Emotional Literacy Assessment and Intervention Ages 7 to 11*. London: NFER Nelson.

Feshbach, N. (1975) Empathy in Children: Some Theoretical and Empirical Considerations. *The Counseling Psychologist*, 5, 2, 25-30.

Fox, M. (2002) The education of children with special educational needs: Evidence or value driven? *Educational and Child Psychology*, 19, 3, 42-53.

Frederickson, N. (2002) Evidence-based practice and educational psychology. *Educational and Child Psychology*, 19, 3, 96-111.

Frederickson, N. and Cline, T. (2002) *Special Educational Needs, Inclusion and Diversity: A Textbook*. Buckingham: Open University Press.

Goleman, D. (1995) *Emotional Intelligence: Why it can matter more than IQ*. London: Bloomsbury Publishing.

Goodman, R. (1997) The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, 38, 5, 581-586.

Goodman, R. (2001) Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 11, 1337-1345.

Goodman, R. and Scott, S. (1999) Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is Small Beautiful? *Journal of Abnormal Child Psychology*, 27, 1, 17-24.

Hallam, S. (2009) An evaluation of the Social and Emotional Aspects of Learning (SEAL) programme: promoting positive behaviour, effective learning and well-being in primary school children. *Oxford Review of Education*, 35, 3, 313-330.

Hallam, S., Rhamie, J. and Shaw, J. (2006) *Evaluation of the Primary Behaviour and Attendance Pilot*. Nottingham: DfES Publications.

Humphrey, N., Kalamvouka, A., Bolton, J., Lendrum, A., Wigelsworth, M., Lennie, C. and Farrell, P. (2008) *Primary Social and Emotional Aspects of Learning (SEAL): Evaluation of Small Group Work*. Nottingham: DCSF Publications.

Humphrey, N., Lendrum, A., Wigelsworth, M. and Kalamvouka, A. (2009) Implementation of primary Social and Emotional Aspects of Learning small group work: a qualitative study. *Pastoral Care in Education*, 27, 3, 219-239.

Joseph, G. and Strain, P. (2003) Comprehensive Evidence-Based Social—Emotional Curricula for Young Children: An Analysis of Efficacious Adoption Potential. *Topics in Early Childhood Special Education*, 2, 62-73.

Kelly, B., Longbottom, J., Potts, F., and Williamson, J. (2004) Applying Emotional Intelligence: Exploring the Promoting Alternative Thinking Strategies Curriculum. *Educational Psychology in Practice*, 20, 3, 221-240.

Kendall, P. (2000) Guiding Theory for Therapy with Children and Adolescents. In Kendall, P. (Ed.) *Child and Adolescent Therapy: Cognitive-Behavioral Procedures (2nd Edition)*. London: The Guilford Press.

Kimber, B. Sandell, R. and Bremberg, S. (2008) Social and Emotional Training in Swedish Schools for the Promotion of Mental Health: an effectiveness study of 5 years of intervention. *Health Education Research*, 23, 6, 931-940.

Lendrum, A., Humphrey, N., Kalambouka, A. and Wigelsworth, M. (2009) Implementing primary Social and Emotional Aspects of Learning (SEAL) small group interventions: recommendations for practitioners. *Emotional and Behavioural Difficulties*, 14, 3, 229-238.

Maddern, L., Franey, J., McLaughlin, V. and Cox, S. (2004) An Evaluation of the Impact of an Inter-Agency Intervention Programme to Promote Social Skills in Primary School Children. *Educational Psychology in Practice*, 20, 2, 135-155.

Maines, B. and Robinson, G. (1988) *B/G-Steem: a Self-esteem Scale with Locus of Control items*. Bristol: Lucky Duck Publishing.

Martin, C., Ruble, D. and Szkrybalo, J. (2002) Cognitive Theories of Early Gender Development. *Psychological Bulletin*, 128, 6, 903-933.

Matthews, B. (2004) Promoting emotional literacy, equity and interest in science lessons for 11-14 year olds; the 'Improving Science and Emotional Development' project. *International Journal of Science Education*, 26, 3, 281-308.

Matthews, G., Roberts, R. and Zeidner, M. (2004a) Seven Myths About Emotional Intelligence. *Psychological Inquiry*, 15, 3, 179-196.

Matthews, G., Zeidner, M. And Roberts, R. (2004b) *Emotional Intelligence: Science and Myth*. London: MIT Press.

Mayer, J. and Cobb, C. (2000) Educational Policy on Emotional Intelligence: Does It Make Sense? *Educational Psychology Review*, 12, 2, 163-183.

Mayer, J., Caruso, D. and Salovey, P. (2000a) Selecting a Measure of Emotional Intelligence: The Case for Ability Scales. In Bar-On, R. and Parker, J. (Eds.) *The Handbook of Emotional Intelligence: Theory, Development, Assessment and Application at Home, School and in the Workplace*. San Francisco: Jossey-Bass.

Mayer, J., Salovey, P. and Caruso, D. (2000b) Emotional Intelligence as Zeitgeist, as Personality, and as a Mental Ability. In Bar-On, R. and Parker, J. (Eds.) *The Handbook of Emotional Intelligence: Theory, Development, Assessment and Application at Home, School and in the Workplace*. San Francisco: Jossey-Bass.

Mertens, D. (1998) *Research Methods in Education and Psychology: Integrating Diversity with Quantitative and Qualitative Approaches*. London: SAGE Publications.

Mertens, D. and McLaughlin, J. (2004) *Research and Evaluation Methods in Special Education*. London: SAGE Publications.

Miller, A. and Todd, A. (2002) Educational psychology and difficult behaviour in schools: Conceptual and methodological challenges for an evidence-based profession. *Educational and Child Psychology*, 19, 3, 82-95.

Novaco, R. (1975) *Anger Control: The Development and Evaluation of an Experimental Treatment*. London: Lexington Books.

Ofsted (2010) *The Evaluation Schedule for Schools*. Retrieved online May 2010 from <http://www.ofsted.gov.uk/Ofsted-home/Forms-and-guidance/Browse-all-by/Other/General/Evaluation-schedule-of-judgements-for-schools-inspected-under-section-five-of-the-Education-Act-2005-from-September-2009>

Pallant, J. (2007) *SPSS Survival Manual (3rd Edition)*. Maidenhead: Open University Press.

Parton, C. and Manby, M. (2009) The contribution of group work programmes to early intervention and improving children's emotional wellbeing. *Pastoral Care in Education*, 27, 1, 5-19.

Perneger, T. V. (1998) What's wrong with Bonferroni adjustments. *British Medical Journal*, 316, 1236-1238.

Pössel, P., Horn, A., Groen, G. and Hautzinger, M. (2004) School-based prevention of depressive symptoms in adolescents: a 6-month follow-up. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 1003-1010.

Qualter, P., Gardner, K., and Whiteley, H. (2007) Emotional Intelligence: Review of Research and Educational Implications. *Pastoral Care*, 25, 1, 11-20.

Qualter, P., Whiteley, H., Hutchinson, J. and Pope, D. (2007) Supporting the Development of Emotional Intelligence Competencies to Ease the Transition from Primary to High School. *Educational Psychology in Practice*, 23, 1, 79-95.

Reichardt, C. (1979) The Statistical Analysis of Data from Nonequivalent Group Designs. In Cook, T. and Campbell, D. (Eds.) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. London: Houghton Mifflin.

Robson, C. (2002) *Real World Research: A resource for social scientists and practitioner-researchers (2nd Edition)*. Oxford: Blackwell Publishing.

Salovey, P. and Mayer, J. (1990) Emotional Intelligence. *Imagination, Cognition and Personality*, 9, 3, 185-211.

Sharp, P. (2001) *Nurturing Emotional Literacy: A Practical Guide for Teachers, Parents, and those in the Caring Professions*. London: David Fulton Publishers.

Sharp, P. and Herrick, E. (2000) Promoting emotional literacy: Anger Management Groups. In Barwick, N. (Ed.) *Clinical Counselling in Schools*. London: Routledge.

Smith, P., O'Donnell, L., Easton, C. and Rudd, P. (2007) *Secondary Social, Emotional and Behavioural Skills (SEBS) Pilot Evaluation*. Retrieved online July 2009 from <http://www.dcsf.gov.uk/research/data/uploadfiles/DCSF-RR003.pdf>

Spencer, C. and Cuthill, A. (2007) *Primary National Strategy Behaviour and Attendance Associate Pilot: Final Report of Associate Pilot September 2004 – August 2006*. Unpublished report for Bedfordshire County Council.

Spivack, G. and Shure, M. (1974) *Social Adjustment of Young Children: A Cognitive Approach to Solving Real-Life Problems*. London: Jossey-Bass.

Tennant, R., Goens, C., Barlow, J., Day, C. and Stewart-Brown, S. (2007) A Systematic Review of Reviews of Interventions to Promote Mental Health and

Prevent Mental Health Problems in Children and Young People. *Journal of Public Mental Health*, 6, 1, 25-32.

Tew, M. (2007) *School Effectiveness: Supporting Success through Emotional Literacy*. London: Paul Chapman Publishing.

Waterhouse, L. (2006) Multiple Intelligences, the Mozart Effect, and Emotional Intelligence: A Critical Review. *Educational Psychologist*, 41, 4, 207-225.

Weare, K. (2004) *Developing the Emotionally Literate School*. London: Paul Chapman Publishing.

Weare, K. and Gray, G. (2003) *What Works in Developing Children's Emotional and Social Competence and Wellbeing?* Nottingham: DfES Publications.

Wells, J., Barlow, J. and Stewart-Brown, S. (2003) A Systematic Review of Universal Approaches to Mental Health Promotion in Schools. *Health Education*, 103, 4, 197-220.

Zeidner, M., Roberts, R. and Matthews, G. (2002) Can Emotional Intelligence Be Schooled? A Critical Review. *Educational Psychologist*, 37, 4, 215-231.

Zins, J. and Elias, M. (2007) Social and Emotional Learning: Promoting the Development of All Students. *Journal of Educational and Psychological Consultation*, 17, 2&3, 233-255.

Zins, J., Elias, M. and Greenberg, M. (2003) Facilitating Success in School and in Life Through Social and Emotional Learning. *Perspectives in Education*, 21, 4, 55-67.

SECONDARY REFERENCES

Achenbach, T., McConaughy, S. and Howell, C. (1987) Child/Adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213-232. Cited in Goodman, R. (2001) Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 11, 1337-1345.

Wolpert, M., Fonagy, P., Belsky, J. and Humphrey, N. (ongoing) *Targeted Mental Health in Schools Evaluation*. Project funded by the Department for Children, Schools and Families, 2008-2011. Cited in Humphrey, N., Kalamvouka, A., Bolton, J., Lendrum, A., Wigelsworth, M., Lennie, C. and Farrell, P. (2008) *Primary Social and Emotional Aspects of Learning (SEAL) Evaluation of Small Group Work*. Annesley: DCSF Publications.

APPENDIX A – STRENGTHS AND DIFFICULTIES QUESTIONNAIRE

Strengths and Difficulties Questionnaire

For each item, please mark the box for Not True, Somewhat True or Certainly True. It would help us if you answered all items as best you can even if you are not absolutely certain or the item seems daft! Please give your answers on the basis of the child's behaviour over the last six months or this school year.

Child's Name.....

Male/Female

Date of birth.....

	Not True	Somewhat True	Certainly True
Considerate of other people's feelings	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Restless, overactive, cannot stay still for long	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often complains of headaches, stomach-aches or sickness	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Shares readily with other children (treats, toys, pencils etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often has temper tantrums or hot tempers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Rather solitary, tends to play alone	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Generally obedient, usually does what adults request	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Many worries, often seems worried	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Helpful if someone is hurt, upset or feeling ill	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Constantly fidgeting or squirming	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Has at least one good friend	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often fights with other children or bullies them	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often unhappy, down-hearted or tearful	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Generally liked by other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Easily distracted, concentration wanders	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Nervous or clingy in new situations, easily loses confidence	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kind to younger children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often lies or cheats	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Picked on or bullied by other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Often volunteers to help others (parents, teachers, other children)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Thinks things out before acting	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Steals from home, school or elsewhere	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Gets on better with adults than with other children	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Many fears, easily scared	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sees tasks through to the end, good attention span	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Signature.....

Date.....

Parent/Teacher/Other (please specify:)

Thank you very much for your help

© Robert Goodman, 2005

APPENDIX B – B/G STEEM

<p>B/G - Steem Primary Scale for Boys</p> <p>Please answer all the questions. Put a ring around YES or NO</p>			
Name	Age	School	Date
1. Is your school work good?		yes no	
2. Do you like being a boy?		yes no	
3. Are you strong and healthy?		yes no	
4. Does someone else always choose what you wear?		yes no	
5. Do your parents think you behave well?		yes no	
6. Do children like playing with you?		yes no	
7. Are you very nice looking?		yes no	
8. Are you as clever as other children?		yes no	
9. Does the teacher notice when you work hard?		yes no	
10. Are you a fast runner?		yes no	
11. Can you make your work better if you really try?		yes no	
12. Are you a good reader?		yes no	
13. Are you good at looking after yourself?		yes no	
14. Does your mum or dad like you to help them?		yes no	
15. Do you choose your friends?		yes no	
16. Do you have a best friend?		yes no	
17. Is your teacher pleased with your work?		yes no	
18. Do you need a lot of help?		yes no	
19. Are your parents usually fair?		yes no	
20. Do you often get the blame when it is not your fault?		yes no	
21. Do you find sums hard?		yes no	
22. Do you have nice clothes?		yes no	
23. Do other people decide everything about your life?		yes no	
24. Are you the best looking in your class?		yes no	
25. Are your parents proud of you?		yes no	
26. Do you think that wishing can make nice things happen?		yes no	
27. Would you like to be someone else?		yes no	

APPENDIX C – QUESTIONNAIRE

Implementation of Primary SEAL in lower schools – a brief survey

Name of school:

I am researching Primary SEAL in Central Bedfordshire and would like to get a picture of practice across the local authority. This survey should take around **3 minutes** to complete. I should be grateful if you would complete the survey and return it to me at the address shown on page two by **6th February 2009** at the latest. Please circle the answers that apply to you and write answers on the dotted lines where necessary.

Name of person completing survey (optional):

Designation of person completing survey (optional):

ALL

1. Has your school adopted the *whole school* SEAL initiative?

Yes (please go to question 2)

No (please go to question 3)

Only if you answered yes to question 1:

2. Does your school have support from the local authority with this?

(eg training from the School Improvement, Behaviour Support or Educational Psychology teams)

Yes (please go to question 4)

No (please go to question 4)

Only if you answered no to question 1:

3. Have you adopted an alternative to the *whole school* SEAL initiative?

Yes Please state what this is.....

No Is there a reason for this?.....

ALL

4. Does your school currently run *small group* SEAL (the silver set materials) as a 'wave two' intervention?

Yes (please go to question 5)

No (please go to question 6)

Only if you answered yes to question 4:

5. Does your school receive support from the Behaviour Support Team in running these groups?

Yes

No

ALL

6. Does your school plan to run *small group* SEAL (the silver set materials) as a 'wave two' intervention in the near future (ie summer term 2009 or autumn term 2009)?

Yes – summer term and autumn term 2009 (please go to your final question, 8)

Yes – summer term 2009 (please go to your final question, 8)

Yes – autumn term 2009 (please go to your final question, 8)

Don't know (please go to your final question, 8)

No (please go to your final question, 7)

Only if you answered no to question 6:

7. Does your school plan to run *an alternative* to small group SEAL (the silver set materials) as a 'wave two' intervention in the near future?

Yes Please state what this is.....

No Is there a reason for this?.....

Only if you answered yes or don't know to question 6:

8. Would your school be interested in taking part in a research project evaluating the impact of small group SEAL (the silver set materials)?

This would involve running two short-term groups for targeted children. I will do some questionnaires with the children before and after the group intervention and supply a report of findings and recommendations to the school. Teachers and parents will be asked to complete the Strengths and Difficulties Questionnaire (25 questions with a tick box response) on two or three occasions during the project.

Yes

Not sure

No

Many thanks for completing this survey. Please return your completed form to me: Clare Otter, Trainee Educational Psychologist, Psychology and Advisory Support Team, County Hall, Cauldwell Street, Bedford, MK42 9AP. Please contact me on (01234) 228693 or clare.otter@bedscc.gov.uk if you have any questions.

APPENDIX D – INFORMATION PACK



Small group SEAL research project – What do I need to know?

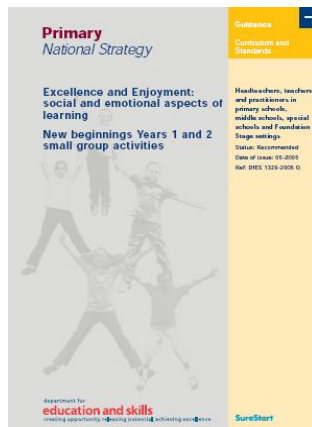


Why is this research project taking place?

The training for educational psychologists has changed to a three-year doctorate course. Trainee educational psychologists spend years 2 and 3 of the course in local authorities and undertake a research project during this time. I have chosen to study the effects of small group SEAL on children in lower schools in Central Bedfordshire as part of my doctorate at the University of Nottingham.

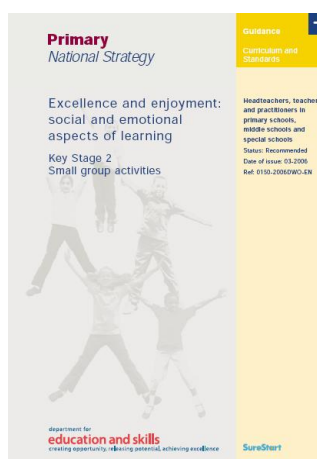
What is small group SEAL? Who is it for? What does it involve?

Small group SEAL is a 'wave two' intervention for children who need extra help to develop social, emotional and behavioural skills, having already taken part in whole school and whole class SEAL work. The government has developed materials to help schools support these pupils. These are called the 'silver set' materials and can be found online.



DfES (2005) Excellence and Enjoyment: Social and Emotional Aspects of Learning. New beginnings Years 1 and 2 small group activities. (Also available for other SEAL themes)

http://nationalstrategies.standards.dcsf.gov.uk/node/65949?uc=force_uj



DfES (2006) Excellence and Enjoyment: Social and Emotional Aspects of Learning. Key stage 2 small group activities.

http://nationalstrategies.standards.dcsf.gov.uk/node/89188?uc=force_uj



Small group SEAL research project – What do I need to know?



What are we committing to if we decide to take part in the project?

Firstly, there are the usual things that schools do when they run small groups:

- Using procedures to select which children should take part in the groups (further guidance on this is in the pack – see pages 4 and 5)
- Contacting parents and gaining their consent for their child to take part in the group (a sample letter is included in the pack – see page 6)
- Providing a member of staff to plan and run the groups and liaise with others
- Providing a room where the group can meet every week and be undisturbed

Then there are additional things that taking part in this project would involve:

1. Pre, post and follow up measures

- Asking class teachers to fill in questionnaires about the pupils' emotional literacy before the group work starts, after the group work has finished and half a term later. The questionnaires are short and involve multiple choice questions. Sample questionnaires are included in this pack – see pages 7-8.
- Asking parents to fill in questionnaires about the pupils' emotional literacy before the group work starts, after the group work has finished and again half a term later. The questionnaires are short and involve multiple choice questions. Sample questionnaires are included in this pack – see pages 9-10.
- Allowing me to come into school and do two questionnaires with the pupils before the group work starts, after the group work has finished and again half a term later. This will involve seeing each pupil outside of the classroom for approximately 10 to 15 minutes.

2. Running two groups so comparisons can be made

- Running two groups consecutively, preferably one group each half term. Don't worry if you think you won't have enough children for two groups, as groups are made up of targeted children and role models (see page 4). Having two groups also means you can give more thought as to which children would or would not work well together in a group.

3. Allowing me to visit the groups

- Allowing me to come and watch one of the sessions and talk to the group facilitator about it. These visits are intended to be supportive and to find out how the groups are organised.

4. Allowing me to use the results

- Allowing me to use the questionnaire results in my research project. Of course, your right to confidentiality and anonymity will be guaranteed. The results will simply be grouped together with all of the other results from schools across the local authority.



Small group SEAL research project – What do I need to know?



What will we get in return? What are the benefits?

I hope that you find your school benefits from taking part in the project.

- ✓ This is a chance to evaluate the effectiveness of your provision for pupils who need extra support. I will feedback data from parents, teachers and the pupils themselves on how the small group work has affected the emotional literacy of the children. This will be in the form of a written summary, which I will be happy to discuss further with you if you would like.
- ✓ The results of the study may lead to improvements in practice for future groups. Depending on the results, I may include some suggestions for future group work in the written report.
- ✓ You will be supported in your practice through feedback and advice if you need it. This support will be on top of your usual time allocation from the Psychology and Advisory Support Team, which means that you will not lose out on the usual amount of time you have from your educational psychologist in a year.

What happens if our circumstances change?

Schools are busy places and the capacity to provide small group support for pupils may vary at different times of the year for several reasons, for example staffing levels, inspections, end of year tests and other events. I fully understand that unforeseen circumstances may affect you and would like to remind you of your right to withdraw from the project at any stage and for any reason.

What if I still have questions about it? How can I find out more?

Please contact me if there is anything else you want to know. My details are as follows:

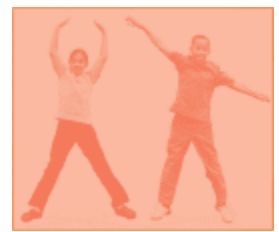
Clare Otter, Trainee Educational Psychologist

Telephone: 01234 228693

Email: clare.otter@centralbedfordshire.gov.uk



Selecting children to take part in the wave two SEAL intervention 1



Targeted children

Children chosen for wave two interventions should be *those who would benefit from early intervention in developing their social, emotional and behavioural skills*. The groups should build on wave one (quality first teaching) work done by the whole school and class by *allowing selected children to practice and extend their learning*. If you think a child's needs require individual intervention supported by professionals external to the school, that child would probably be better supported at wave three level. Government guidance suggests that *groups of children from the same class* may work best.

The *selection process for including children in a group should be transparent*. You may like to use the criteria below to help your decision-making:

- A Children who have got an Individual Education Plan with social, emotional or behaviour targets, an Individual Behaviour Plan or a Pastoral Support Plan*
- B Children who have been permanently excluded from their previous school
- C Children who have had fixed term exclusions in the past term
- D Children who have negative entries about their behaviour in the County Behaviour Log at least once a week*
- E Children who have been assessed as having difficulties with social and emotional development according to assessments such as the Boxall Profile, the Strengths and Difficulties Questionnaire, the Emotional Literacy Assessment Instrument or another standardised test or questionnaire*
- F Children who have joined the school or class within the last half term
- G Children who are 'left out' by other children (eg. identified by a sociogram)*
- H Children who have recently experienced considerable change to their life (eg. a child who is finding things difficult following parental separation, the birth of a sibling, being taken into care, moving house, the death of a relative)
- I If you use other ways to identify children for wave two SEAL, let me know)

**If you need advice and support with the techniques listed above, please contact me*

Children chosen to model social, emotional and behavioural skills

Groups should be balanced and include children with a range of needs. They should include children who can act as role models for the skills being taught in the group.

Other group composition considerations

You may also want to get a balance of boys and girls and of different personalities.

The research project

Schools taking part in the project will be required to run two consecutive groups. This design means that school staff can plan for their ideal group composition in advance. For example, if there were two children who struggle to manage frustration you might want to put one in each group.



Selecting children to take part in the wave two SEAL intervention 2



Name of school

People involved in the selection process could be the:

- Class Teacher
- Group Facilitator
- Co-facilitator (if applicable)
- SEAL/Behaviour and Attendance co-ordinator (or someone with a similar 'whole school' responsibility)

Use the information in 'Selecting children to take part in the wave two SEAL intervention 1' to choose members of your two groups. Remember that your reasons for choosing the children should be clear and that groups should have a good mix of needs and skills.

You should aim to include about six children in each group.

Name of child	Targeted child or role model?	Reason for inclusion in the group (you can write the letter from sheet 1)	Group 1 or group 2?



Small group SEAL sample letter to parents



Please feel free to use this sample letter by photocopying it onto school headed paper. Alternatively, you could use it as an idea to get you started and write your own letter.

Dear Parent/Carer,

Your child has been chosen to take part in some small group work in school. These small groups are a fun way for children to practice their social and emotional skills.

We want to see how much the groups help children's skills at school and at home. You can help us do this by filling in some questionnaires before and after the group work.

I hope you will agree for your child to take part in the group. If so, please fill in the reply slip below and return it to school.

Your sincerely,

.....

Please fill in this reply slip and return it to by
.....

Child's name

- ☒ I agree for my child to take part in the small group work
- ☒ I agree to fill in the questionnaires before and after the group work

Signed (Parent/Carer)

Date

APPENDIX E – SHEET FOR GATHERING INFORMATION ON PARTICIPANTS

Name of school:

Children's details

Name of pupil	Year/Class	Date of birth	Intervention before/after half term	Targeted (and reason for inclusion)/role model	Gender	Ethnicity	EAL? (If so, first language?)	SEN stage?

Group leader(s) details

Name:

Professional role:

Gender:

School details

Years since taking on wave two SEAL?:

Support received in setting up wave two SEAL?:

APPENDIX F – TOOL FOR ASSESSING INTERVENTION INTEGRITY

Small group SEAL session checklist

(Based on the DCSF silver set materials)

Name of school:
Name of group leader:
Name of observer:
Date of observation:

Observation during the session

Criterion	✓ or ✗	Notes
Is there a welcome /check in activity where children can air feelings or concerns?		
Is there at least one warm up activity for example a circle game or round?		
Is there a reminder of group aims/rules for appropriate behaviour?		
Is work from previous sessions referred to or progress in the past week discussed?		
Is the plan for the session/learning outcome shared with the children?		
Is there a core activity ?		
Is there a chance for the children to review and reflect on the session? For example they may rate how well they have met the learning outcome.		
Is there a task for the children to carry out before the next session?		
Is there a relaxation activity , for example quiet time or visualisation?		
Is the session at least 40 minutes long ?		

Discussion with group leader before or after the session

Criterion	✓ or ✗	Notes
Do you have access to any supervision arrangements ? For example someone to speak to about your experience of the groups or advice on how to deal with issues that arise?		
Do you have a chance to meet with the class teacher(s) weekly to review the session and plan the next one?		
Are the children working on the same SEAL theme in class as they are in the small group?		
Will there be at least six sessions ?		
Does the session take place in a room which is free from interruptions ?		
Is children's progress in the group shared with parents/carers ?		
Were the children told why they were chosen for the group?		