Evaluating the impact of whole-class self-management and interdependent group contingency approaches on pupil engagement and disruptive behaviour

Kamal Bhana

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

July 2017

Word count: *41,561*

Abstract

This study investigates the efficacy of two whole-class approaches to classroom management, *self-management* and *interdependent group contingency,* in a sample of 8-9 year olds in the UK.

Phase A investigates which approach is most effective in reducing off-task and disruptive behaviours in target lessons, and in improving behaviour in general. Phase B investigates whether combining the approaches further reduces off-task and disruptive behaviour, and improves general behaviour. The research employed a quasi-experimental design. In Phase A, pupils were allocated to one of four conditions: self-management (n=30), interdependent group contingency (n=29), waitlist control receiving daily rule reminders (n=28), or a waitlist control who continued as usual (n=26). The approaches were delivered by class teachers over four-weeks. In Phase B, the class receiving self-management in Phase A, received interdependent group contingency as well, for a further four weeks. The waitlist control group continued as per Phase A.

Pre- and post-test measures for both phases were obtained through structured observations of whole-class on-task, off-task and disruptive behaviours. Teachers also completed the *Strengths and Difficulties Questionnaire* (SDQ) for each pupil. Findings indicated that self-management and interdependent group contingency reduced off-task behaviour, however only interdependent group contingency reduced disruptive behaviour. Combining the approaches led to no further reductions in these behaviours. SDQ data suggested that self-management, either alone or combined with interdependent group contingency, had no significant impact on general behaviour. However, interdependent group contingency alone, appeared to lead to greater general behaviour concerns.

The findings are reviewed in light of the literature with limitations acknowledged. Avenues for future research are also identified. In conclusion, this research presents tentative evidence supporting the efficacy of these individual approaches for off-task and/or disruptive behaviour. Findings that the combined approach is not efficacious and

that neither approach improves general behaviour, should be interpreted cautiously given the study's limitations.

Acknowledgements

"It always seems impossible until it's done."

Nelson Mandela

It is with deepest gratitude that I express appreciation to the many people who supported and encouraged me on this journey through the research project and the Doctorate as a whole:

- To the tutor team at The University of Nottingham, with particular thanks to Nathan Lambert for his encouragement and guidance throughout the course.
- To the Educational Psychologists in the Local Authority in which the research was conducted, for their assistance with the research project. In particular, thank you to my placement supervisor and the psychology assistants, without whom this particular research project may not have been possible.
- To all of the staff in the school that participated in this research. Thank you to the management team for welcoming the idea of participating and in particular, thank you to the teachers who implemented the approaches and supported data collection.
- To the Trainee Educational Psychologists in Cohort 9 who have been a most valuable source of guidance and encouragement over the past three years.
- To my mother, Nayan, my father, Narendra, and to my sisters, Jessica and Dina, who always believed in me and offered help in any way they could.
- And finally, to my fiancé, Pritul, for his unwavering support, endless patience and positivity throughout this journey.

Contents

List of tables xiv

List of figures xvi

List of acronyms and abbreviations xviii

1. INTRODUCTION 1

1.1. Introduction 1

1.2. The researcher's personal and professional interest 2

1.3. Structure of the thesis 3

2. LITERATURE REVIEW 4

2.1. Introduction 4

2.2. Behaviour in schools 4

2.2.1. Pupil outcomes 5

2.3. Psychological theories of behaviour and engagement 6

2.3.1. Instrumental conditioning 6

2.3.2. Social interdependence theory 7

2.3.3. Self-determination theory 8

2.3.4. Self-regulation theory 10

2.4. Approaches to classroom management 12

2.4.1. A history of classroom management 12

2.4.2. Components of effective classroom management 13

2.4.3. Classroom management programmes 13

2.4.4. Implementation science 14

2.4.5. Interdependent group contingency 16

2.4.5.1. Aims and process 16

2.4.5.2. Evidence base 16

2.4.5.3. Strengths and limitations 17

2.4.6. Self-management 19
2.4.6.1. Aims and process 19
2.4.6.2. Evidence base 20
2.4.6.3. Strengths and limitations 21
2.4.6.4. Using contingent reinforcement 22
2.4.7. Summary 23
2.5. Systematic review 23
2.5.1. Aims 24
2.5.2. Criteria for study selection 24
2.5.3. Search strategy and identification of studies 26
2.5.4. Assessment of study quality 27
2.5.5. Review of studies 27
2.5.5.1. Self-management and interdependent group contingency combined 31
2.5.5.2. Impact of self-management or interdependent group contingency separately and combined 31
2.5.6. Approach procedures, data collection methods and data analysis 35
2.5.7. Limitations of the review 36
2.5.8. Summary of the review 36
2.6. Research questions 38
3. METHODOLOGY 39
3.1. Real world research 39
3.2. Evidence-based practice 40
3.3. Philosophical paradigms in research 41
3.3.1. Positivism and post-positivism 41
3.3.2. Constructivism 42
3.3.3. Pragmatism 43
3.4. Research designs 43

3.4.1. Fixed designs 43
3.4.1.1. Quasi-experimental designs 44
3.4.2. Flexible designs 45
3.4.3. Mixed method designs 45
3.5. The current research 45
3.5.1. The paradigm 45
3.5.2. The design of the current Study 46
3.5.3. Stakeholders 49
3.5.3.1. University of Nottingham 49
3.5.3.1. The researcher 49
3.5.3.2. The Local Authority 49
3.5.3.3. School staff, parents and pupils 49
3.5.3.3.1. Engagement 49
3.5.4. School and community context 51
3.5.5. Sampling procedure 51
3.5.6. Participants 52
3.5.6.1. Classes 52
3.5.7. Implementing the approaches 53
3.5.7.1. The procedures 53
3.5.7.2. Implementation 56
3.5.7.3. Teacher and pupil training 56
3.5.7.4. Fidelity checks 57
3.5.7.5. Timeline for the research 57
3.5.8. Measures 58
3.5.8.1. Strengths and Difficulties Questionnaire 58
3.5.8.2. Observation schedules 59
3.5.8.2.1. Piloting for inter-observer agreement 59
3.5.8.2.2. Categorising disruptive and on- and off-task behaviour 60

3.5.8.2.3. Observation procedures 61

3.5.8.2.4. Analysis 62

3.5.8.2.5. Inter-observer agreement checks during the study 62

3.5.9. Issues of data quality 63

3.5.9.1. Reliability 63

3.5.9.2. Internal validity 64

3.5.9.3. Threats to external validity 67

3.5.10.Ethical considerations 67

3.5.10.1. Autonomy and dignity of persons 68

3.5.10.2. Maximising benefit and minimising harm and risk 69

4. RESULTS 71

4.1. Introduction 71

4.2. Approach to data analysis 73

4.2.1. Statistical tests 73

4.2.1.1. Power 74

4.2.2. Visual inspection 75

4.2.3. Analyses chosen for this research 77

4.2.3.1. Visual inspection of observation data 77

4.2.3.1.1. Inter-rater reliability checks of graphed data 77

4.2.3.2. Statistical analysis of Strengths and Difficulties Questionnaire data 78

4.2.3.2.1. Descriptive statistics 78

4.2.3.2.2. Non-parametric tests and significance levels 81

4.3. Phase A results 82

4.3.1. Observation data 83

4.3.1.1. Class 1 (self-management) 84

4.3.1.1.1. Implementation of the approach 84

4.3.1.1.2. Off-task behaviour 84

4.3.1.1.3. Disruptive behaviour 86

4.3.1.2. Class 2 (interdependent group contingency) 88
4.3.1.2.1. Implementation of the approach 88
4.3.1.2.2. Off-task behaviour 88
4.3.1.2.3. Summary of findings: Off-task behaviour 90
4.3.1.2.4. Disruptive behaviour 90
4.3.1.2.5. Summary of findings: Disruptive behaviour 92
4.3.1.3. Class 3 (rule reminder) 92
4.3.1.3.1. Implementation of the approach 92
4.3.1.3.2. Off-task behaviour 92
4.3.1.3.3. Disruptive behaviour 95
4.3.1.4. Class 4 (no change) 96
4.3.1.4.1. Implementation of the approaches 96
4.3.1.4.2. Off-task behaviour 97
4.3.1.4.3. Disruptive behaviour 98
4.3.2. Combined data 100
4.3.3. Results summary of observation data for research question 1 103
4.3.4. Results summary of observation data for research question 2 103
4.3.5. Results summary of observation data for research question 3 104
4.3.6. Strengths and Difficulties Questionnaire data 104
4.3.6.1. Behavioural problems 104
4.3.6.2. Hyperactivity and concentration problems 106
4.3.6.3. Results summary of Strengths and Difficulties Questionnaire data for research question 1 109
4.3.6.4. Results summary of Strengths and Difficulties Questionnaire data for research question 2 109
4.3.6.5. Results summary of Strengths and Difficulties Questionnaire data for research question 3 110

4.3.7. Overall summary of results for Phase A 110
4.4. Phase B results 111
4.4.1. Observation data 111
4.4.1.1. Class 1 (interdependent group contingency added to self-management) 111
4.4.1.1.1. Implementation of the approach 111
4.4.1.1.2. Off-task behaviour 112
4.4.1.1.3. Disruptive behaviour 113
4.4.1.2. Class 3 (rule reminder) 115
4.4.1.2.1. Implementation of the approach 115
4.4.1.2.2. Off-task behaviour 115
4.4.1.2.3. Disruptive behaviour 117
4.4.1.3. Class 4 (no change) 118
4.4.1.3.1. Implementation of the approaches 118
4.4.1.3.2. Off-task behaviour 118
4.4.1.3.3. Disruptive behaviour 120
4.4.2. Strengths and Difficulties Questionnaire data 124
4.4.2.1. Results on the behavioural problems subscale 124
4.4.2.2. Results on the hyperactivity and concentration problems subscale 126
4.4.3. Overall summary of results for Phase B 127
5. DISCUSSION 128
5.1. Introduction 128
5.2. Summary of findings 128
5.2.1. Aims of Phase A 128
5.2.2. Research question 1 128
5.2.2.1. Hypotheses 129
5.2.2.2. Key findings and links to literature 129

5.2.2.3. Alternative interpretations and limitations of this research question 130
5.2.2.4. Conclusion 131
5.2.3. Research question 2 131
5.2.3.1. Hypotheses 131
5.2.3.2. Key findings and links to literature 132
5.2.3.3. Alternative interpretations and limitations of the research 133
5.2.3.4. Conclusion 134
5.2.4. Research question 3 134
5.2.4.1. Hypotheses 134
5.2.4.2. Key findings and links to literature 135
5.2.4.3. Alternative interpretations and limitations of the research 136
5.2.4.4. Conclusion 136
5.2.5. Phase A: Further links to the literature 137
5.2.6. Aims of Phase B 138
5.2.7. Research question 4 139
5.2.7.1. Hypotheses 139
5.2.7.2. Key findings and links to literature 140
5.2.7.3. Alternative interpretations 141
5.2.7.4. Conclusion 142
5.3. A wider focus on factors affecting engagement and disruption 142
5.4. Limitations of the research 144
5.4.1. Research design 144
5.4.2. Behaviour of the classes 147
5.4.3. Measures 148
5.4.3.1. Observation schedule 148
5.4.3.2. Strengths and Difficulties Questionnaire 149

5.4.4. Data analysis 150
5.4.5. The approaches 152
5.4.6. Implementation and fidelity 152
5.4.7. Inter-observer agreement 157
5.4.8. Extraneous variables associated with real world research in schools 158
5.5. Implications of the findings 160
5.5.1. Implications for schools 160
5.5.2. Implications for Educational Psychologists 161
5.6. Future research 162
5.7. Researcher reflections 165
6. CONCLUSION 167
6.1. Introduction 167
6.2. Unique contribution of this research 167
6.3. Key findings 168
7. REFERENCES 170
8. APPENDICES 197
8.1. Appendix: Search strategy 198
8.2. Appendix: Screening flow diagram 200
8.3. Appendix: Study characteristics 201
8.4. Appendix: Weight of evidence ratings 207
8.5. Appendix: Ethical approval letters 210
8.6. Appendix: Stakeholder engagement letters 214
8.7. Appendix: Head teacher and Chair of Governers' research information sheet and consent form 217
8.8. Appendix: Teachers' research information sheet and consent form 222
8.9. Appendix: Parent information sheets and consent forms 227
8.10. Appendix: Approach procedures 237

8.11. Appendix: Teacher training slides for Phase A of the study 240
8.12. Appendix: Teacher training slides for Phase B of the study..... 247
8.13. Appendix: Resources used for the approaches 255
8.14. Appendix: Teacher record sheet 258
8.15. Appendix: Fidelity checklists 259
8.16. Appendix: Outcomes of the fidelity checks..... 263
8.17. Appendix: Strengths and Difficulties Questionnaire..... 265
8.18. Appendix: Observation schedule 267
8.19. Appendix: Debrief letter..... 268
8.20. Appendix: Timeline of Research Process 269
8.21. Appendix: Alpha level correction details 271
8.22. Appendix: Inter-rater reliability check sheet..... 274

List of tables

Table 2.1: A table outlining the five main factors that affect the successful implementation of an intervention, as outlined by Damschroder et al., (2009). 15
Table 2.2: A table to show the inclusion and exclusion criteria used for study selection in the systematic review. 26
Table 2.3: A summary of the studies selected for the systematic review.. 30
Table 3.1: A table to illustrate the design of the current study. 48
Table 3.2: A table outlining the contextual information of the classes included in this study. 53
Table 3.3: A table to outline the procedures for the classroom approaches used in this study. 55
Table 3.4: A table to outline the threats to reliability in this study and steps taken to reduce them. 64
Table 3.5: A table to outline the threats to internal validity in this study and steps taken to reduce them. 66
Table 4.1: A table to present the descriptive statistics of the SDQ data. ... 80
Table 4.2: A table to outline the findings from the visual inspection, on Class 1's graphed off-task behaviour data at Time 1 and Time 2. 85
Table 4.3: A table to outline the findings from the visual inspection, on Class 1's graphed disruptive behaviour data at Time 1 and Time 2. .. 87
Table 4.4: A table to outline the findings from the visual inspection, on Class 2's graphed off-task behaviour data at Time 1 and Time 2. 89
Table 4.5: A table to outline the findings from the visual inspection, on Class 2's graphed disruptive behaviour data at Time 1 and Time 2. .. 91
Table 4.6: A table to outline the findings from the visual inspection, on Class 3's graphed off-task behaviour data at Time 1 and Time 2. 93
Table 4.7: A table to outline the findings from the visual inspection, on Class 3's graphed disruptive behaviour data at Time 1 and Time 2. .. 96
Table 4.8: A table to outline the findings from the visual inspection, on Class 4's graphed off-task behaviour data at Time 1 and Time 2. 98
Table 4.9: A graph to show the percentage of disruptive intervals observed in Class 4, at Time 1 and Time 2. 99

Table 4.10: A table to outline the findings from the visual inspection, on Class 1's graphed off-task behaviour data at Time 2 and Time 3..... 113
Table 4.11: A graph to show the percentage of disruptive intervals observed in Class 1, at Time 2 and Time 3..... 114
Table 4.12: A table to outline the findings from the visual inspection, on Class 3's graphed off-task behaviour data at Time 2 and Time 3..... 116
Table 4.13: A graph to show the percentage of disruptive intervals observed in Class 3, at Time 2 and Time 3..... 118
Table 4.14: A table to outline the findings from the visual inspection, on Class 4's graphed off-task behaviour data at Time 2 and Time 3..... 119
Table 4.15: A graph to show the percentage of disruptive intervals observed in Class 4, at Time 2 and Time 3..... 121
Table 5.1: A table outlining the factors that may have affected implementation of the approaches in this research, as well as ideas for remediation, based on the five areas that affect successful implementation, as outlined by Damschroder et al., (2009)..... 156
Table 8.1: A table to outline the databases and search terms used to locate journal articles in the systematic review, as well as the numbers of papers included and excluded..... 199
Table 8.2: A table to summaries the characteristics of the studies reviewed in the systematic review..... 206
Table 8.3: A table to show the Weight of Evidence A judgements for the systematic review papers..... 207
Table 8.4: Weight of Evidence ratings for the reviewed papers..... 209
Table 8.5: A table to outline the detailed procedures for the approaches used in this research..... 239
Table 8.6: A table to summarise the outcomes from the fidelity checks conducted in this research. 264
Table 8.7: A table to show which statistical tests yielded statistically significant results according to Holm's (1979) correction and as a result of employing a p=.01 level of significance..... 273

List of figures

Figure 2.1: A reproduction of Zimmerman's (2000) model of self-regulation 11
Figure 2.2: A diagram to show the stages involved in conducting a systematic review (Gough, 2007)..... 24
Figure 2.3: Descriptors for Gough's (2007) Weight of Evidence Framework. 27
Figure 3.1: An example of a hierarchy of evidence from health, adapted from Concato, Shah and Horwitz (2000). 40
Figure 4.1: A graph to show the percentage of off-task intervals observed in Class 1, at Time 1 and Time 2. 85
Figure 4.2: A graph to show the percentage of disruptive intervals observed in Class 1, at Time 1 and Time 2. 87
Figure 4.3: A graph to show the percentage of off-task intervals observed in Class 2, at Time 1 and Time 2. 89
Figure 4.4: A graph to show the percentage of disruptive intervals observed in Class 2, at Time 1 and Time 2. 91
Figure 4.5: A graph to show the percentage of off-task intervals observed in Class 3, at Time 1 and Time 2. 93
Figure 4.6: A graph to show the percentage of off-task intervals observed in Class 3, at Time 1 and Time 2, with the anomalous data point removed. 94
Figure 4.7: A graph to show the percentage of disruptive intervals observed in Class 3, at Time 1 and Time 2. 95
Figure 4.8: A graph to show the percentage of off-task intervals observed in Class 4, at Time 1 and Time 2. 97
Figure 4.9: A graph to show the percentage of disruptive intervals observed in Class 4, at Time 1 and Time 2. 99
Figure 4.10: A graph to show the percentage of off-task intervals observed in all classes, at Time 1 and Time 2. 101
Figure 4.11: A graph to show the percentage of disruptive intervals observed in all classes, at Time 1 and Time 2. 102

Figure 4.12: A graph to show all classes' median scores on the behavioural problems subscale of the SDQ, at Time 1 and Time 2. 105
Figure 4.13: A graph to show all classes' median scores on the hyperactivity and concentration problems subscale of the SDQ, at Time 1 and Time 2. 106
Figure 4.14: A boxplot to illustrate each class' spread of data on the hyperactivity and concentration problems subscale at Time 1. 108
Figure 4.15: A boxplot to illustrate each class' spread of data on the hyperactivity and concentration problems subscale at Time 2. 108
Figure 4.16: A graph to show the percentage of off-task intervals observed in Class 1, at Time 2 and Time 3. 112
Figure 4.17: A graph to show the percentage of disruptive intervals observed in Class 1, at Time 2 and Time 3. 114
Figure 4.18: A graph to show the percentage of off-task intervals observed in Class 3, at Time 2 and Time 3. 116
Figure 4.19: A graph to show the percentage of disruptive intervals observed in Class 3, at Time 2 and Time 3. 117
Figure 4.20: A graph to show the percentage of off-task intervals observed in Class 4, at Time 2 and Time 3. 119
Figure 4.21: A graph to show the percentage of disruptive intervals observed in Class 4, at Time 2 and Time 3. 120
Figure 4.22: A graph to show the percentage of off-task intervals observed in classes 1, 3 and 4 at Time 2 and Time 3 122
Figure 4.23: A graph to show the percentage of disruptive intervals observed in classes 1, 3 and 4 at Time 2 and Time 3 123
Figure 4.24: A graph to show class 1, 3 and 4's median scores on the behavioural problems subscale of the SDQ, at Time 2 and Time 3. .. 125
Figure 4.25: A graph to show class 1, 3 and 4's median scores on the hyperactivity and concentration problems subscale of the SDQ, at Time 2 and Time 3. 126

List of acronyms and abbreviations

EP	Educational Psychologist
IGC	Interdependent group contingency
IOA	Inter-observer agreement
NC	No change
PA	Psychology Assistant
RCT	Randomised control trials
RR	Rule reminder
SDQ	Strengths and Difficulties Questionnaire
SM	Self-management

1. Introduction

1.1. **Introduction**

The current study aims to answer the question: *How effective is a whole-class self-management approach, and how effective is an interdependent group contingency approach in improving overall behaviour in junior school classes?*

Self-management, as a whole class approach, involves pupils monitoring, recording, and analysing their own behaviour (Davies & Witte, 2000). The aim is to develop pupils' skills in self-regulation (Freeman & Dexter-Mazza, 2004; King-Sears, 2008; Niesyn, 2009; Rooney, Hallahan, & Lloyd, 1984; Traxson, 1994).

Interdependent group contingency requires all pupils to follow the rules in order to earn a whole-class reward (McKissick, Hawkins, Lentz, Hailley, & McGuire, 2010; C. H. Skinner, 2004). The aim is to encourage engagement through reward contingencies (B. F. Skinner, 1953) and through developing positive social interdependence in which pupils conform to expectations because their behaviour will impact the outcomes of their peers as well as themselves (Deutsch, 1949; D W Johnson & Johnson, 2005).

This study focuses on evaluating the impact of these two approaches on off-task and disruptive behaviours in the classroom.

Behaviour in schools is of on-going concern in the UK (Department for Education, 2017; Department for Education and Science, 1989; Department for Education and Skills, 2005; Haydn, 2012; Wheldall & Merrett, 1984), contributing to teacher stress (Axup & Gersch, 2008; Gu & Day, 2013) and to poorer outcomes for young people (Finn & Zimmer, 2012; Froiland & Oros, 2014; Ladd & Dinella, 2009). Whilst these approaches have been shown to be effective in other countries (Chafouleas, Sanetti, Jaffery, & Fallon, 2012; Denune et al., 2015; Ennis, Blair, & George, 2016; Hoff & Ervin, 2013; Murphy, Theodore, Aloiso, Alric-Edwards, & Hughes, 2007) no research on self-management and interdependent group contingency approaches appears to have been conducted with a UK population.

The current research also, crucially, adopts an experimental group design; an approach as yet under-represented in the research literature (Bruhn, McDaniel, & Kreigh, 2015). The study also aims to overcome some of the limitations of previous research, such as floor, ceiling, and ordering effects (Denune et al., 2015) and a lack of standardised measurement and statistical analyses.

Finally, there is at present limited research considering the efficacy of self-management as a whole-class approach, and self-management paired with interdependent group contingency. As such, the current study aims to contribute towards an evidence base for these classroom management approaches.

1.2. **The researcher's personal and professional interest**

The researcher first developed an interest in investigating the factors that affect the behaviour of children and young people in classrooms during her time as a primary school teacher. Having observed a number of classroom management approaches implemented by teachers with varying degrees of success, the researcher decided to make 'engagement' the focus of her Master's dissertation, for which a case study methodology was employed.

More recently, in her role as a Trainee Educational Psychologist the researcher has conducted various classroom observations. These have led to a renewed interest in classroom management approaches, and a particular interest in investigating which approaches might be most effective in promoting engagement.

After investigating the literature, research suggesting the efficacy of interdependent group contingency and self-management was discovered. However, gaps in the research literature considering these approaches were also apparent. The intention of the current research is to address some of these gaps.

1.3. **Structure of the thesis**

Chapter 2: Literature Review *– This chapter reviews the literature on engagement and disruptive behaviour in schools including outcomes for pupils and teachers. Psychological theories and models are reviewed as well as the evidence-base for self-management and interdependent group contingency approaches. Avenues for further research are identified and research questions are developed.*

Chapter 3: Methodology *– This chapter discusses paradigms and issues surrounding real world research before outlining possible methodologies. The design and methodology for this study are described and justified. Finally, ethical issues and issues of data quality are considered.*

Chapter 4: Results *– This chapter discusses and justifies the approaches taken for data analysis and presents the results.*

Chapter 5: Discussion *– This chapter presents a brief summary of the findings, linking this to key literature and discussing alternative interpretations in light of the study's limitations. Implications of the findings for schools and Educational Psychologists are outlined and directions for future research are identified.*

Chapter 6: Conclusion *– This chapter outlines the key findings from the research in relation to the research questions and highlights the unique contributions made to the literature.*

2. Literature Review

2.1. Introduction

The aim of this literature review is to explore and discuss the relevant literature surrounding the present area of study; the issue of behaviour in schools and classroom management approaches.

The review will first outline the issue of behaviour in UK schools and highlight the impact of disengagement and disruption on teachers and pupils. The perspectives on behaviour of four psychological theories will be outlined and discussed; these are instrumental conditioning, social interdependence theory, self-regulation theory and self-determination theory. The review will briefly outline the history of classroom management, the components of effective classroom management as highlighted in the literature, offer a brief overview of classroom management approaches used in schools and discuss in detail two approaches; self-management and interdependent group contingency. Finally, a systematic review evaluating the current evidence base for the effectiveness of whole-class self-management and interdependent group contingency on classroom behaviour will be presented.

The author will argue that the current evidence base for the whole-class use of these two approaches is limited, and that the research which does exist is flawed. Finally, the research questions for the current study will be outlined.

2.2. Behaviour in schools

Concerns around behaviour in UK classrooms have been reported for decades (Department for Education, 2017; Department for Education and Science, 1989; Department for Education and Skills, 2005; Haydn, 2012; Wheldall & Merrett, 1984), with teachers feeling ill-prepared to manage (Briesch, Briesch, & Chafouleas, 2015), and many considering leaving the profession as a result (Association of Teachers and Lecturers, 2010; Department for Education, 2012).

Disruptive behaviours have been defined as including being out of seat, engaging in vocalisations without permission, distracting others by making noises (Denune et al., 2015), violating rules and disrupting the learning of others by throwing objects (Hoff & Ervin, 2013). Although these behaviours are challenging for teachers (Clunies-Ross, Little, & Kienhuis, 2008; Hulac & Benson, 2010), so is a lack of engagement such as not attending to the task (Ofsted, 2005).

Engagement is considered to be a multifaceted construct (Christenson, Reschly, & Wylie, 2012; Fredricks, Blumenfeld, & Paris, 2004) in which behavioural (following expectations and completing work), emotional (positive or negative feelings which affect eagerness to engage) and cognitive engagement (motivation, the use of learning strategies and exertion of mental effort) are interrelated (Fredricks et al., 2004), and inseparable from motivation (Reeve, 2012). It is argued that engagement is the observable form of motivation (Reeve, 2012), and both are considered to be influenced by the context/environment (Reeve, 2012; Reschly & Christenson, 2012). Motivation is considered necessary, but insufficient alone, for engagement to occur (Appleton, Christenson, & Furlong, 2008).

2.2.1. **Pupil outcomes**

It is recognised that disruption reduces learning opportunities for all when teachers respond to these behaviours (Clunies-Ross et al., 2008; Gordon, 2001; Mitchem et al., 2001; Murphy et al., 2007; Oliver, Wehby, & Reschly, 2011; Stage & Quiroz, 1997), which is likely to negatively impact the attainment of all pupils (Finn & Zimmer, 2012). In extreme cases, this may lead to exclusions (Department for Education, 2014), which have been linked to feelings of shame, resentment, (Partington, 2001), isolation and depression (Leyden & Miller, 1998; C. Wright, Weekes, & McGlaughlin, 2000).

Disengagement can also lead to poorer outcomes. Research has found that high levels of disengagement predicted low attainment in later school life (Froiland & Oros, 2014; Ladd & Dinella, 2009; Rowe & Rowe, 1992), whereas high engagement led to greater feelings of competence and

connection to the school community (E. A. Skinner & Pitzer, 2012), and supported later work competence in adulthood (Masten et al., 2010).

2.3. Psychological theories of behaviour and engagement

A number of psychological theories have been proposed to explain engagement and disruptive behaviours (Ayers, Clarke, & Murray, 2000; Frederickson & Cline, 2015; Porter, 2000). The theories that are most pertinent to this research and underpin the classroom management approaches under investigation will be discussed below.

2.3.1. Instrumental conditioning

Rooted in the behaviourist perspective, instrumental (or operant) conditioning views behaviour as learned through interaction with the environment (Blackman, 1984; B. F. Skinner, 1938, 1953). Behaviour is considered to be repeated more frequently as a result of its expression being positively reinforced (Gleitman, Fridlund, & Reisberg, 2004; B. F. Skinner, 1953). Where a behaviour leads to no reinforcement or consequences which are considered punishing, those behaviours will be engaged in less frequently (Gleitman et al., 2004). According to this perspective, disruptive behaviour in the classroom may occur frequently due to those behaviours being reinforced, perhaps through positive peer attention (Altman & Linton, 1971; Northup et al., 1995), whereas engaging with tasks may be less reinforcing or even punishing.

Early studies investigated the use of praise (McAllister, Stachowiak, Baer, & Conderman, 1968), vicarious reinforcement of other pupils, and token economy systems (Altman & Linton, 1971; Osborne, 1968), on disruptive behaviours in classrooms. These studies indicated that such strategies could be effective in increasing engagement (Altman & Linton, 1971; Bednar, Zelhart, Greathouse, & Weinberg, 1970; Burchard & Tyler, 1964; McAllister et al., 1968; Osborne, 1968). Despite much research in favour of the behaviourist perspective, it has been criticised for ignoring the importance and influence of cognitions and emotions on motivation and behaviour (Nicolson & Ayers, 2004), for not addressing the underlying

reason for the behaviour's occurrence (Mufti & Peace, 2012; Trevithick, 2011), and for ignoring historical factors which may be contributing to the behaviour (Cross, 2004).

2.3.2. **Social interdependence theory**

Social interdependence theory was originally postulated by Deutsch (1949), extending the work of Lewin (1935) who stated that when an individual perceives a desired goal, tension is created, which motivates the individual to work towards that goal. Social interdependence refers to the way in which an individual's goals are related to those of others (D W Johnson & Johnson, 1994). Interdependence can be positive, where an individual is able to meet their goal if others also do, or negative, where an individual can only meet their goal if others fail to (Deutsch, 1949; D W Johnson & Johnson, 2005). According to this theory, if interdependence is positive, the following psychological processes will be created: *substitutability,* (the degree to which an individual's actions may be substituted for another's), *inducibility,* (the degree of openness to influencing others and being influenced by them), and *positive cathexis* (the amount of positive psychological energy one invests on external objects) (Deutsch, 1949; D W Johnson & Johnson, 2008). These processes are hypothesised to lead to individuals engaging in more co-operative behaviours to support their peers.

Social interdependence theory states that disruptive behaviour may occur where teachers' and pupils' goals are incompatible (D W Johnson & Johnson, 2006). A pupil must first be motivated to achieve a goal (Lewin, 1935), such as staying on task to receive a reward. In this way, the teacher and pupils' goals align. Social interdependence is believed to create responsibility among group members, motivating individuals to work towards a joint goal, because one does not want to fail others (D W Johnson, 2003).

In support of this theory, D. W. Johnson, Maruyama, Johnson, Nelson and Skon (1981) found that co-operative goal structures were superior to competition and individualistic efforts in promoting achievement and productivity. Furthermore, research indicates that classroom behaviour

improves when a class can only receive a reward if all pupils follow expectations (Ennis et al., 2016; Hartman & Gresham, 2016; Kelshaw-Levering, Sterling-Turner, Henry, & Skinner, 2000; Ling & Barnett, 2013; McKissick et al., 2010; Murphy et al., 2007; Theodore, Bray, & Kehle, 2004).

D. W. Johnson and Johnson (2005) outlined limitations to Deutsch's original theory. For instance, the theory assumes that there are no power differences between individuals which may affect individuals' behaviour. Also the theory assumes that past experiences would not impact upon whether one behaves positively or negatively in a social interdependence situation.

Instrumental conditioning and social interdependence theory underpin interdependent group contingency, which aims to encourage engagement and reduce disruption through positive reinforcement and through fostering social interdependence.

2.3.3. **Self-determination theory**

Self-determination theory (SDT) was developed by Deci and Ryan (1985) to explain human motivation. This is relevant to the study of classroom engagement as it is argued that motivation and engagement are inextricably linked (Reeve, 2012). To be self-determined means to problem solve, goal set, self-manage and develop a positive self-image (Clark, Olympia, Jensen, Heathfield, & Jenson, 2004), which are key factors for learning and engagement.

SDT states that all humans are innately motivated to have their need for autonomy, competence and relatedness met (Deci & Ryan, 1985; Reeve, 2012; Ryan & Deci, 2000). Autonomy describes being in control of one's own behaviour and acting in one's own interest, relatedness describes the need to connect with others and competence describes feeling capable in achieving at a task (Deci & Ryan, 1985; Ryan & Deci, 2002). When these needs are met, humans experience high intrinsic motivation which increases their optimal functioning and engagement (Corpus & Wormington, 2014; Deci & Ryan, 1985; Hayenga & Corpus, 2010; Lemos & Veríssimo, 2014; Ryan & Deci, 2000; Schultz & Switzky, 1993; Taylor et al., 2014). Intrinsic motivation is the desire to do an activity for the personal

pleasure or satisfaction that it brings, whereas extrinsic motivation is the desire to engage in an activity in order to obtain an external outcome or reward (Niemiec & Ryan, 2009; Ryan & Deci, 2000).

Sociocultural influences can present a barrier to having these needs met (Deci & Ryan, 1985; Reeve, 2012; Ryan & Deci, 2000). This is thought to lead to a reduction in intrinsic motivation and engagement (Oga-Baldwin, Nakata, Parker, & Ryan, 2017; Reeve, 2012; Standage, Duda, & Ntoumanis, 2005). Research has highlighted that the following factors can affect engagement:

- School climate (Fredricks et al., 2004).
- Communications around effort and ability (Reschly & Christenson, 2012), such as negative (Koka & Hagger, 2010), or positive feedback (Deci, Koestner, & Ryan, 1999; Hattie & Timperley, 2007; Koka & Hagger, 2010).
- Choice and co-operative endeavours (Fredricks et al., 2004; Reeve, Deci, & Ryan, 2004).
- Class size (Blatchford, Bassett, & Brown, 2011).
- Teacher autonomy-, competence- and relatedness-support (Garon-Carrier et al., 2015; Jang, Reeve, & Deci, 2010; Klem & Connell, 2004; Lavigne, Vallerand, & Miquelon, 2007; F. Mitchell, Gray, & Inchley, 2015; Shih, 2008; Sparks, Dimmock, Whipp, Lonsdale, & Jackson, 2015; Standage et al., 2005; Stroet, Opdenakker, & Minnaert, 2013; Van den Berghe, Cardon, Tallir, Kirk, & Haerens, 2016).
- Peer exclusion or victimisation (Buhs, Ladd, & Herald, 2006).
- Classroom structures (Fredricks et al., 2004).
- Task characteristics (Fredricks et al., 2004).
- The degree to which these factors support or undermine autonomy, relatedness and competence (Reschly & Christenson, 2012).

Despite the wealth of research into SDT, it appears that the majority of the studies focus on physical education activities with adolescents (F. Mitchell et al., 2015; Sparks et al., 2015; Standage et al., 2005; Van den Berghe et al., 2016), or with international samples (Burton, Lydon, D'alessandro, & Koestner, 2006; Cheon & Reeve, 2015; Jang et al., 2010; Nie & Lau, 2009;

Oga-Baldwin et al., 2017; Shih, 2008; Taylor et al., 2014; Van den Berghe et al., 2016). There appears a dearth of research exploring SDT in primary-aged or British samples, with academic tasks. This makes generalisability of these findings difficult. Furthermore, many of these studies relied on interview data (Sparks et al., 2015) or self-report measures (Burton et al., 2006; Oga-Baldwin et al., 2017; Shih, 2008; Standage et al., 2005) rather than direct observations of engagement, which could have introduced bias.

2.3.4. **Self-regulation theory**

Whilst SDT gives a broad outline of the factors which impact motivation and thereby engagement, self-regulation theory (SRT) focuses on the internal processes by which an individual is able to regulate their thoughts, feelings and behaviours. Zimmerman (2000, p. 14) defined self-regulation as "self-generated thoughts, feelings and actions that are planned and cyclically adapted to the attainment of personal goals". A number of theories and models of self-regulation have been developed in different fields (Boekaerts, Pintrich, & Zeidner, 2000).

Social-cognitive self-regulation theories such as those proposed by Bandura (1991) and Zimmerman (2000), hypothesise that behaviour is regulated through reflecting on its causes and consequences (Bandura, 1991) through feedback from previous experiences (Zimmerman, 2000). Both viewed the regulation of behaviour as involving forethought, which involves the setting of proximal and distal goals, and contains self-motivation beliefs such as self-efficacy, outcome expectations and intrinsic interest (Bandura, 1991; Zimmerman, 2000). In Zimmerman's (2000) theory, self-regulation also involves performance or volitional control which consists of self-observation and self-control towards the goal, and self-reflection which involves making judgements and self-evaluations of one's performance towards one's own goals. Figure 2.1 illustrates Zimmerman's model.

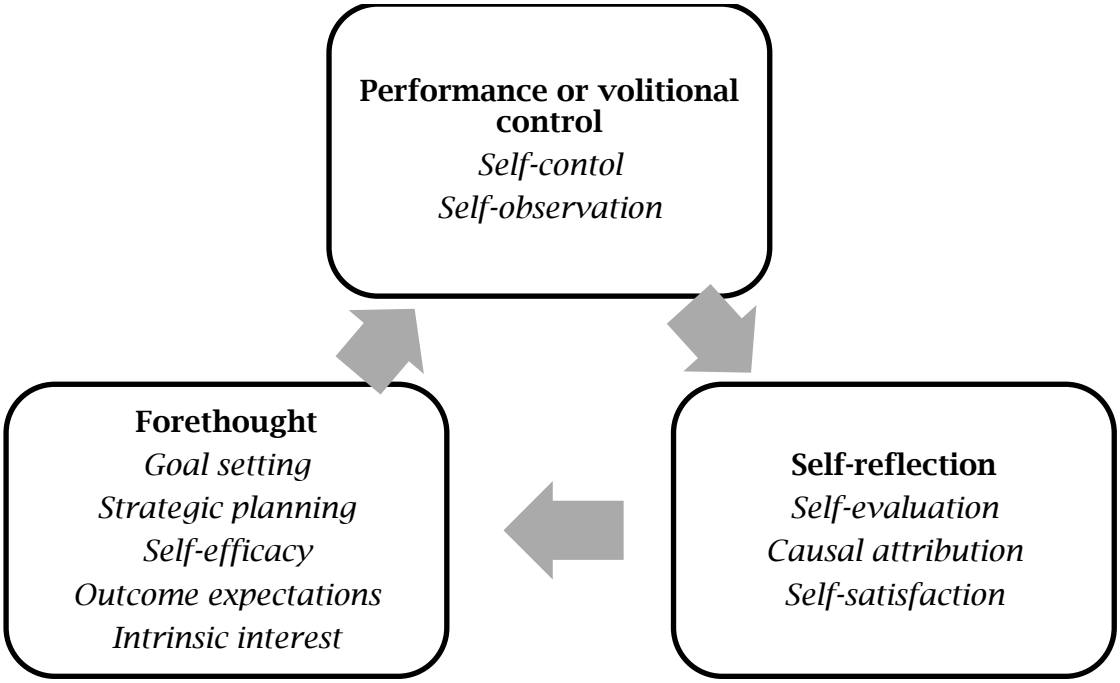


Figure 2.1: A reproduction of Zimmerman's (2000) model of self-regulation

Overall, social-cognitive theories of self-regulation state that whether a pupil engages in the classroom depends on their thoughts around whether they can achieve at the task (self-efficacy), whether they will achieve (outcome expectations) based on feedback, and social-environmental factors such as how their friends are performing and whether the task is communicated as being difficult (Urdan & Schoenfelder, 2006). All of these factors are thought to affect an individual's motivation to engage with classroom tasks (Urdan & Schoenfelder, 2006).

In support of these theories, a recent study by Lee, Lee and Bong (2014) found that self-efficacy and personal interest predicted academic self-regulation in a South Korean sample. These findings were also supported by O'Keefe and Linnenbrink-Garcia (2014). Furthermore, studies around physical activity in schools have found that some components of social cognitive theories such as self-efficacy and social support explained some of the variance in physical activity behaviour, but not all (Dewar et al., 2013; J. Martin, McCaughtry, Flory, Murphy, & Wisdom, 2011; Ramirez, Kulinna, & Cothran, 2012). Research also suggested that support from

peers and teachers motivated pupils to follow classroom rules and to be co-operative (Wentzel, 1997, 1998).

These findings and theories highlight the complexities surrounding the wide range of factors which affect engagement in class, and in turn, inform the type of interventions to be used to overcome disengagement. Self-determination theory and self-regulation theory underpin self-management which aims to support pupils in becoming more autonomous, competent and intrinsically motivated to follow expectations, by teaching self-regulation through self-observation and self-evaluation.

2.4. Approaches to classroom management

Classroom management has been defined as "the actions teachers take to create an environment that supports and facilitates both academic and social-emotional learning" (Evertson & Weinstein, 2006, p. 4).

2.4.1. A history of classroom management

Over the last 100 years, classroom management research has progressed from focusing behaviourally on the use of routines, rewards and sanctions in the early 20th century (Bagley, 1907), to focusing on teacher effectiveness (Brown, 1966; Flanders, 1961; Nelson, 1963; Willower, 1960), and classroom climate on pupil learning and behaviour (G. J. Anderson, Walberg, & Welch, 1969; Brophy, 2006; Lewin, Lippitt, & White, 1939; McDill, Meyers Jr, & Rigsby, 1967). However the influence of behaviourism through the use of rules, praise and ignoring continued into the 1960s and is still popular today (Madsen Jr, Becker, & Thomas, 1968; O'Leary, Becker, Evans, & Saudargas, 1969; Ward, 1971; Yawkey, 1971). In the 1960s, the view that pupils could be involved in developing rules and be responsible for managing their own behaviour began to emerge (Lovitt & Curtiss, 1969; Wehlitz, 1960) and in the 1970s, researchers looked towards social and cognitive explanations of, and interventions for behaviour (Bandura, 1986, 1991; Brophy, 2006; Meichenbaum, 1977), such as the importance of peer influence (Solomon & Wahler, 1973), self-regulation skills (Carter & Doyle, 2006; Lovitt & Curtiss, 1969; McLaughlin, 1976), and maladaptive thoughts

(Meichenbaum, 1977). More recently the impact of the physical environment (Barrett, Davies, Zhang, & Barrett, 2015; Blatchford et al., 2011; Brophy, 2006; Stronge, Tucker, & Hindman, 2004), children's social and emotional skills (Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Pech, 2013), and the quality of the teacher-pupil relationship (Evertson & Weinstein, 2006; Hajdukova, Hornby, & Cushman, 2014; Marsh, 2012), on behaviour have been considered.

2.4.2. **Components of effective classroom management**

Hart (2010) and Marzano (2003) identified a number of features of effective classroom management. These were the use of rules (Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008; Stronge et al., 2004), positive reinforcement of appropriate behaviours (Simonsen et al., 2008), responding to undesired behaviours, positive staff-pupil interactions (O'Connor, Dearing, & Collins, 2011), high teacher expectations in lessons (Stronge et al., 2004; Urhahne, 2015), agreed procedures for managing misbehaviour (Hart, 2010; Marzano, 2003), and a positive classroom environment (M. M. Mitchell & Bradshaw, 2013; Stronge et al., 2004). Research indicates that positive approaches are more effective (Herrera & Little, 2005; M. M. Mitchell & Bradshaw, 2013), whereas reactive strategies are associated with higher teacher stress and decreased on-task behaviour (Clunies-Ross et al., 2008). These findings have important implications for what components classroom management strategies should include and interdependent group contingency and self-management aim to promote some of these.

2.4.3. **Classroom management programmes**

A number of classroom programmes have been developed from psychological theory. The foci of these approaches are variable and can include an emphasis on internal behaviour control, external behaviour control, classroom ecology, interpersonal relationships, curriculum and discourse (Evertson & Weinstein, 2006; Wubbels, 2011). Although over time, perspectives on classroom behaviour have included cognitive, interpersonal and ecological aspects (Ayers et al., 2000), the use of reward charts, token economy systems, giving praise (Harlacher, 2015; Wheldall &

Merrett, 1984), and the use of sanctions (Department for Education, 2016; Department for Education and Skills, 2012) remain popular. *The Good Behaviour Game* is a classroom management approach which uses the behaviourist principles of reward to increase appropriate behaviour through the contingent reinforcement of a group (Barrish, Saunders, & Wolf, 1969; Flower, McKenna, Bunuan, Muething, & Vega, 2014; Lannie & McCurdy, 2007; Tingstrom, Sterling-Turner, & Wilczynski, 2006). *Assertive Discipline* (Canter, 2010; Canter & Canter, 1976) also uses behaviourist principles.

The present research focuses on two approaches to classroom management; self-management and interdependent group contingency. These are two approaches that have been reported to be effective. They are also notably simple to administer (Bruhn et al., 2015; Hoff & Ervin, 2013; Stage & Quiroz, 1997), which is of particular importance since research has suggested that teachers are more accepting of approaches which are simple and quick to implement (Briesch, Briesch, et al., 2015), over approaches that require more time, greater skill and more effort (Calvert & Johnston, 1990). However, crucially, the research upon which claims of positive effect have been made, was not conducted in the UK. Furthermore, research looking at their *combined* effectiveness is limited.

2.4.4. **Implementation science**

It has long been recognised that despite the growing body of evidence-based interventions which high quality research suggests leads to positive outcomes (Forman et al., 2013), often these interventions are not successfully implemented or maintained in real world settings (Forman et al., 2013). Implementation science is an area of interest in which researchers have focused on exploring the factors that help and hinder the implementation of interventions in the real world (Forman et al., 2013; Kelly, 2012; Olswang & Prelock, 2015). It is particularly important to consider such factors when the intervention is to be implemented by third parties, as in the present research.

Damschroder et al., (2009) developed the 'Consolidated Framework for Implementation Research' based on a range of theories and models of

implementation, in which they concluded that five main factors affect implementation and these are described in Table 2.1.

Factors	Examples
Intervention characteristics	*Stakeholders' perceptions of the intervention's quality, complexity and evidence base.* *The advantage/cost of implementing the intervention compared to alternatives.*
Outer settings	*External and peer pressures, policies or regulations.*
Inner settings	*Communication, norms, values and assumptions (culture) within the organisation.* *The perceived priority of the intervention compared to other priorities.* *Incentives or rewards within the organisation.*
Individual characteristics	*Implementer's attitudes, readiness for change and enthusiasm for the intervention.* *The degree to which the individual identifies with the organisation.*
Process of implementation	*Whether the organisation has been prepared for implementation such as through understanding the needs of the context and planning high quality training.* *Measuring, reflecting and evaluating on whether the intervention is executed correctly.*

Table 2.1: A table outlining the five main factors that affect the successful implementation of an intervention, as outlined by Damschroder et al., (2009).

In support of this model, research highlighted that the following factors affected whether an intervention was successfully implemented: the structure/organisation of the school, intervention characteristics, school policies and values, assistance with training and support from management (Fixsen, Naoom, Blase, Friedman, & Wallace, 2005; Forman & Barakat, 2011; Powell et al., 2012), whether potential barriers to implementation were identified, whether executives were involved, whether relationships were built, whether appropriate training materials were developed, whether implementers were given incentives and whether there were opportunities to shadow and meet with other implementers (Powell et al., 2012). As such it appears that for successful implementation to take place, the wider social context must be considered with organisational support implemented.

2.4.5. **Interdependent group contingency**

2.4.5.1. Aims and process

Interdependent group contingency has been one of the most widely recommended classroom management approaches (Maggin, Johnson, Chafouleas, Ruberto, & Berggren, 2012). It aims to reduce disruptive behaviour and increase engagement through rewarding appropriate behaviour (B. F. Skinner, 1953), promoting positive interdependence (Deutsch, 1949), and by aligning the teacher and pupils' goals (Lewin, 1935). It typically involves specifying a behaviour for the whole class to follow, such as 'allow others to learn', identifying the criterion/goal that the group must reach, such as 'earn 50 points', choosing a reward and administering it to all when the criterion is reached (McKissick et al., 2010; C. H. Skinner, 2004). Rewards are often activities, objects or edible rewards (Maggin et al., 2012).

Interdependent group contingency differs from independent and dependent group contingency (Ennis et al., 2016; Hulac & Benson, 2010; Litow & Pumroy, 1975). In independent group contingency, individuals receive a reward for meeting a goal regardless of whether others in the group also meet the goal (Ennis et al., 2016; Hulac & Benson, 2010; Litow & Pumroy, 1975). In dependent group contingency, the whole group receives a reward based on whether an individual or small group meet the criteria (Ennis et al., 2016; Hulac & Benson, 2010; Litow & Pumroy, 1975). Group contingencies of all kinds require that the target behaviours, criteria and rewards are consistent across all members of the group (Hartman & Gresham, 2016; Kelshaw-Levering et al., 2000; McKissick et al., 2010).

2.4.5.2. Evidence base

There exists a substantial body of evidence to suggest that group contingencies are effective in reducing whole class disruptive behaviours and increasing engagement in both special and mainstream classrooms, with pre-schoolers, primary-aged pupils and adolescents (Christ & Christ, 2006; Ennis et al., 2016; Hansen & Lignugaris-Kraft, 2005; Hartman & Gresham, 2016; Kelshaw-Levering et al., 2000; Ling & Barnett, 2013; Ling,

Hawkins, & Weber, 2011; McKissick et al., 2010; Murphy et al., 2007; Theodore et al., 2004; Williamson, Campbell-Whatley, & Lo, 2009), however, a number of studies in the group contingency literature do not report fidelity checks, which may mean that fidelity issues were present, compromising the validity of those findings (Christ & Christ, 2006; Kelshaw-Levering et al., 2000; Maggin et al., 2012).

Researchers have also sought to establish which group contingency approach is the most effective in reducing disruption and increasing engagement in lessons. To date, the findings have yielded mixed results (Ennis et al., 2016; Hartman & Gresham, 2016; Kelshaw-Levering et al., 2000; Theodore et al., 2004). Ennis et al., (2016) found all group contingencies to be successful in reducing classroom disruption, with evidence of generalisation and maintenance in all classrooms. Hartman and Gresham (2016) found that both dependent and interdependent group contingencies were effective for reducing disruptive behaviour in classrooms, however interdependent group contingency appeared superior over time and these findings were supported by a meta–analysis conducted by Maggin et al., (2012). As such, interdependent group contingency was chosen over the other group contingencies for this research.

2.4.5.3. Strengths and limitations

Studies have suggested that teachers consider interdependent group contingency acceptable, appropriate and manageable to implement (Kelshaw-Levering et al., 2000; Maggin et al., 2012; McKissick et al., 2010; Murphy et al., 2007; Theodore et al., 2004). The advantages of employing group contingencies is that it is more economical and practical to deliver rewards to a whole group, than individually (Chafouleas et al., 2012; Davies & Witte, 2000; Ennis et al., 2016; Hartman & Gresham, 2016; Kelshaw-Levering et al., 2000; Ling et al., 2011), which increases the likelihood that the approach will be conducted to fidelity (Hulac & Benson, 2010; Theodore et al., 2004), and allows multiple pupils and behaviours to be targeted (Maggin et al., 2012). However, it is argued that pupils who do not follow the rules may find themselves coerced or rejected by other members of the group (Hulac & Benson, 2010; Ling et al., 2011). Where

pupils have explicitly been discouraged from behaving negatively towards others and have practised ways to react positively, no rejection or negativity has occurred (Davies & Witte, 2000).

Other potential risks of this approach are that while a reduction in the target behaviour is observed, other inappropriate behaviours are increased because the reward is not contingent upon the presence of those other behaviours (McKissick et al., 2010). Equally, if pupils realise that they will be unable to meet the target in the specified time period, the undesired behaviour may return or the pupils may feel frustrated or disappointed (McKissick et al., 2010). Furthermore, if the chosen rewards are not reinforcing for all of the pupils, this may result in the maintenance of some pupils' undesired behaviours, affecting the group's access to the reward and therefore impact upon the whole class' motivation to engage (Hulac & Benson, 2010; Kelshaw-Levering et al., 2000; C. H. Skinner, Williams, & Neddenriep, 2004). These risks can be reduced through randomisation; choosing rewards at random (Ennis et al., 2016), as well as randomly choosing at the end of the lesson, the behaviour that the class needed to have followed to earn their points (Hulac & Benson, 2010; Kelshaw-Levering et al., 2000; McKissick et al., 2010; Theodore, Bray, Kehle, & Jenson, 2001). Maggin et al., (2012) argued that group contingency interventions may be most effective for short periods such as during a single lesson, as this allows for a thick reinforcement schedule and for rewards to be obtained quickly.

Despite the strong evidence base for group contingencies, its success appears to lie within the receipt of an extrinsic reward. As such, group contingency interventions do not appear to develop intrinsic motivation or the skills to manage one's own behaviour without a reward (Brophy, 2006), suggesting limited efficacy over time. This is illustrated time and again in withdrawal design studies which demonstrate that disruptive or inappropriate behaviours increase once the intervention is removed (Kelshaw-Levering et al., 2000; Ling & Barnett, 2013; Ling et al., 2011; Murphy et al., 2007). As such, group contingencies may not promote the development of pupils' social-emotional growth and self-regulation, which

is considered important for successful classroom management (Evertson & Weinstein, 2006).

2.4.6. **Self-management**

2.4.6.1. Aims and process

Self-management and other terms such as self-monitoring, self-recording and personal goal setting have been used interchangeably in the literature (Barry & Haraway, 2005; Chafouleas et al., 2012). Self-management approaches aim to reduce inappropriate behaviours through explicitly teaching skills in self-monitoring, self-evaluating and self-directing (Freeman & Dexter-Mazza, 2004; King-Sears, 2008; Niesyn, 2009; Rooney et al., 1984; Traxson, 1994). They aim to place the control for behaviour management with the pupil, which increases their autonomy and competence (Deci & Ryan, 1985; Ryan & Deci, 2000), rather than with the teacher, allowing the teacher to teach (Freeman & Dexter-Mazza, 2004; King-Sears, 2008; S.-H. Lee, Simpson, & Shogren, 2007; Mitchem et al., 2001). Through increased autonomy and competence, it is hoped that following rules would become intrinsically motivating, leading to lasting change (Deci & Ryan, 1985; Ryan & Deci, 2000). Self-management also aims to support pupils in regulating their behaviour by increasing their feelings of self-efficacy and outcome expectations of being able to follow the rules (Bandura, 1991; Zimmerman, 2000).

Self-management typically only requires at least two of the following to be present: self-monitoring (an individual observing their behaviour and recording it when cued) (Bruhn et al., 2015; Mitchem et al., 2001), self-evaluation (matching one's self-rating with somebody else's rating; possibly teacher's or a peer's) (Davies & Witte, 2000; Hoff & Ervin, 2013; Mitchem & Young, 2001; Mitchem et al., 2001), and positive reinforcement (Mitchem & Young, 2001). It is argued that teacher feedback is necessary in order for pupils to learn to accurately self-monitor (Freeman & Dexter-Mazza, 2004). In contrast to group contingencies, self-management programmes are reported to teach skills that are maintained over time and generalized to other settings (Freeman & Dexter-Mazza, 2004; Mitchem et al., 2001; Mooney, Ryan, Uhing, Reid, & Epstein, 2005).

2.4.6.2. Evidence base

A wealth of research supports the efficacy of self-management approaches as a targeted intervention for reducing disruptive behaviour and increasing the engagement of individuals and small groups in special education settings, in mainstream primary and secondary schools, and with pupils with ADHD, autism, and learning disabilities (Cavalier, Ferretti, & Hodges, 1997; Dalton, Martella, & Marchand-Martella, 1999; DuPaul, Weyandt, & Janusis, 2011; Freeman & Dexter-Mazza, 2004; Gureasko-Moore, Dupaul, & White, 2006; Holifield, Goodman, Hazelkorn, & Heflin, 2010; Kern, Dunlap, Childs, & Clarke, 1994; King-Sears, 2008; S.-H. Lee et al., 2007; Massey & Wheeler, 2000; Miller, Strain, Boyd, Jarzynka, & McFetridge, 1993; Mitchem et al., 2001; Rafferty, 2012; Rooney et al., 1984; Vance, Gresham, & Dart, 2012; Webber, Scheuermann, McCall, & Coleman, 1993).

Historically, self-management was utilised predominantly in special education settings (Mitchem & Young, 2001). There was then a call for more research to be conducted in mainstream settings (Bruhn et al., 2015; Mooney et al., 2005). Rooney et al., (1984) explored whether self-management procedures could be used with individuals in mainstream settings and reported its efficacy in increasing attention behaviours in four pupils. Miller et al., (1993) implemented self-management as a group intervention in a regular preschool setting and reported improved on-task behaviours and decreased disruptive behaviours in four preschool children.

Recently, it has been suggested that self-management may be effective as a whole-class approach in mainstream settings for improving engagement and behaviour (Hoff & Ervin, 2013; Mitchem & Young, 2001; Niesyn, 2009), as it has been shown to be efficacious in special education settings, when used class-wide (Kern et al., 1994; Salend, Reeder, Katz, & Russell, 1992; Terenzi, Ervin, & Hoff, 2010). Currently, the evidence base for self-management as a whole-class approach to classroom management in mainstream classrooms is limited.

Mitchem, Young, West and Benyo (2001) implemented a peer-assisted self-management programme in three mainstream classrooms with 11-12 year olds, to increase on-task behaviour. They reported that on-task behaviours increased and that these effects were maintained even after the programme had been withdrawn. This finding supports claims made, that self-management programmes develop skills which can be maintained (Freeman & Dexter-Mazza, 2004; Mitchem et al., 2001; Mooney et al., 2005). The teacher also reported increased productivity. However, it is difficult to ascertain whether the maintenance observed after withdrawal of the intervention was due to increased skill in self-management; it may be that as a result of implementing this approach, the teacher's actions and attitudes towards the pupils changed and this is what impacted on-task behaviour and continued to impact it once the approach was removed.

Bruhn et al., (2015) in their meta-analysis of self-management interventions, reviewed 41 studies, in which only one used a group design (Yeo & Choi, 2011). As such, further research on self-management using group designs has been requested (Bruhn et al., 2015).

2.4.6.3. Strengths and limitations

Using self-management in whole-class contexts is an efficient approach that circumvents the time and effort required to administer individualised programmes for several pupils within a classroom (Hoff & Ervin, 2013). It is considered acceptable to teachers and easy to implement (Mitchem & Young, 2001). Furthermore, it is argued that all pupils can benefit from being taught self-management skills (Mitchem & Young, 2001). Studies have suggested that it can be effective to include peer assistance in which peers give each other feedback on their ratings (Mitchem & Young, 2001; Mitchem et al., 2001), as it further empowers pupils to self-monitor accurately, without much of the onus placed on the teacher to provide this feedback.

A major limitation of almost all of the studies on self-management as well as group contingency, is that the research has mainly used single-subject or multiple baseline designs in which the participants serve as their own control (Dalton et al., 1999; Hoff & Ervin, 2013; Holifield et al., 2010;

Kelshaw-Levering et al., 2000; Ling et al., 2011; McKissick et al., 2010; Mitchem & Young, 2001; Murphy et al., 2007; Williamson et al., 2009). As such, the evidence base for both of these approaches rests on the findings from visual analyses. Due to the lack of statistical analysis, it is difficult to ascertain whether any observed improvements are significant. This may increase the risk of a Type 1 error, where the null hypothesis is inaccurately rejected. These studies are also at risk of maturation effects and the likelihood that changes are attributable to environmental factors, which may compromise the validity of the findings (Barlow & Hersen, 1984; Palincsar & Parecki, 1995), due to the lack of a control group.

2.4.6.4. Using contingent reinforcement

With the move towards researching the effectiveness of self-management as a whole-class approach to classroom management, it has been suggested that self-management and group contingency might complement each other, as behavioural improvements from self-management can be maintained and generalised across settings (Freeman & Dexter-Mazza, 2004; Mitchem et al., 2001; Mooney et al., 2005), to account for the issues of maintenance and generalisation with group contingencies (Freeman & Dexter-Mazza, 2004; Maggin et al., 2012; Reiber & McLaughlin, 2004). Furthermore, adding group contingency could make self-management more effective by increasing motivation through the chance to receive a reward (Reiber & McLaughlin, 2004). It is argued that accurate self-monitoring is essential for behavioural improvements to be observed, which may be facilitated through the use of contingent rewards (Reiber & McLaughlin, 2004).

Studies which have investigated the efficacy of self-management and interdependent group contingency alone versus combined, as a *targeted intervention* have yielded mixed results. Bruhn et al., (2015) found that self-management reduced challenging behaviour with and without contingent reinforcement, however others have argued that including reinforcement improves behaviour further (Ardoin & Martens, 2004; Graham-Day, Gardner III, & Hsin, 2010). These studies are few and the sample sizes were very small, making generalizability of the findings difficult. Furthermore, studies which have employed self-management

and/or group contingency have often also included other treatment components, such as function-based support (Bruhn et al., 2015; Vance et al., 2012). As such, it is difficult to draw firm conclusions when reported benefits could be attributable to the additional components.

2.4.7. Summary

Much research has indicated that group contingency (in particular, interdependent group contingency) is effective in reducing whole-class disruptive behaviours and increasing engagement. Self-management also has a strong evidence base as a targeted intervention in these areas, however research on its utility as a whole-class approach, rather than as a targeted intervention, is limited. Research into both interdependent group contingency and self-management lacks scientifically rigorous group designs, calling into question the reliability, validity and generalisability of the findings. Recently, researchers have sought to determine the effectiveness of using self-management and group contingency combined as a whole-class approach to classroom management, as they are each thought to compensate for the shortcomings of the other approach.

A small number of studies have measured the combined impact of a whole-class self-management with interdependent group contingency approach to classroom management and compared their effectiveness separately. These studies are most pertinent to the aims of the current study and therefore will be reviewed and discussed in the next section, in order to establish the current evidence base.

2.5. Systematic review

A systematic review is "a literature review that is designed to *locate, appraise* and *synthesise* the best available evidence relating to a specific research question to provide *informative* and *evidence-based* answers" (Dickson, Cherry, & Boland, 2014, p. 3). The stages involved in a systematic review include (Figure 2.2):

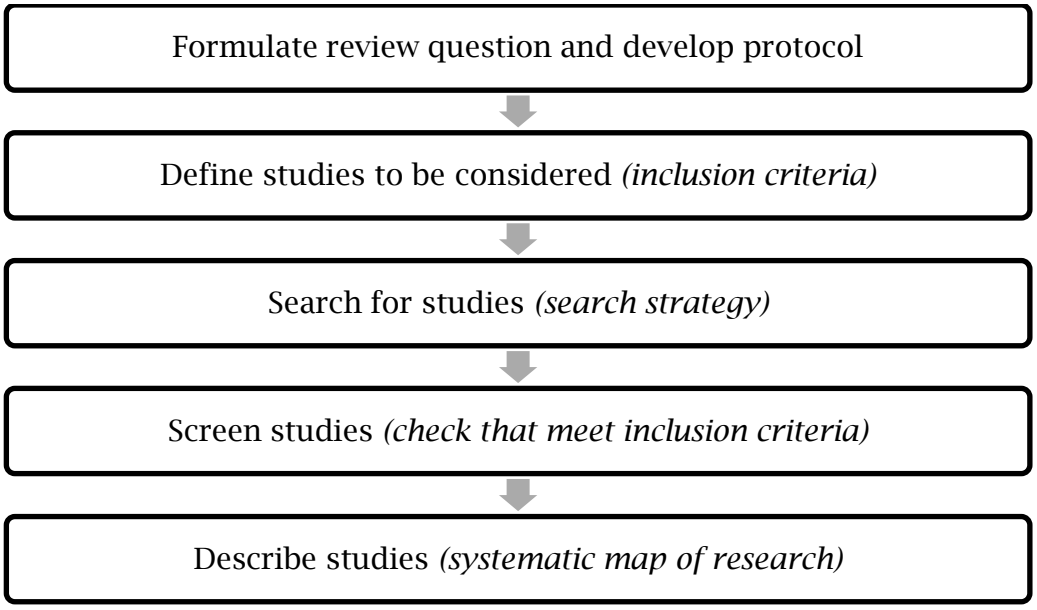


Figure 2.2: A diagram to show the stages involved in conducting a systematic review (Gough, 2007).

The systematic review process also includes appraising the quality and relevance of the studies, synthesising the findings to answer the review question and communicating and engaging with the reader (Gough, 2007).

2.5.1. **Aims**

The aims of this systematic review were to evaluate the impact of combining whole-class self-management with interdependent group contingency approaches on engagement and behaviour in the classroom, and to gain a picture of the strength of its current evidence base. An additional aim was to identify avenues for further research.

2.5.2. **Criteria for study selection**

The following inclusion and exclusion criteria were applied to the searches (Table 2.2):

	Include	Exclude
Population	• Primary-aged children • Adolescents	• Pre-school children • Neonates • Adults • Whole class learning disabilities population
Setting	• School setting • Non-mainstream school setting (e.g. EBD school)	• Pre-school setting • Clinic • Other non-educational setting
Interventions / Approaches	• Whole-class self-management WITH whole-class interdependent group contingency • Peer/teacher feedback element	• Small group/individual self-management • Small group interdependent group contingency • Dependent group contingency • Independent group contingency • Only one of: whole-class self-management OR whole-class interdependent group contingency • Any other additional intervention components
Comparators	• Self-management AND/OR interdependent group contingency + control / waitlist control / comparison group • Self-management AND/OR interdependent group contingency with pre- and post- / repeated measures	• Studies without either a control/comparison group or pre- and post- / repeated measures
Outcomes	• Levels of disruptive behaviour • Levels of engagement • Individuals, small groups or whole class outcomes	• Health or well-being outcomes • Academic attainment outcomes

	Include	Exclude
Study design	• Randomized Control Trials • Quasi-experimental designs • Single-case experimental design • Withdrawal designs • Repeated measures design	• Meta-analyses / Systematic Reviews • Case studies • Other review
Other	• Peer-reviewed • In English	• Non peer-reviewed • In a language other than English

Table 2.2: A table to show the inclusion and exclusion criteria used for study selection in the systematic review.

The inclusion and exclusion criteria were chosen in order ensure that the author obtained the most relevant research and could highlight the current evidence base for the whole-class use of self-management and interdependent group-contingency in non-learning disabled populations.

2.5.3. Search strategy and identification of studies

Key word searches were conducted using *PsycINFO, IngentaConnect, Web of Science, ERIC (EBSCO)* and *Google Scholar* databases, as they contain a number of psychology and education research papers. The following key words were used:

- Self-management AND group contingency *(for PsychINFO, IngentaConnect, Web of Science and ERIC (EBSCO)).*
- "Self-management" AND "interdependent group contingency" AND "class" *(for Google Scholar, as the key words 'self-management AND group contingency' produced 21,100 hits).*

The total number of hits produced was 167. Three additional papers (Glynn, Thomas, & Shee, 1973; T. Johnson, Stoner, & Green, 1996; Mitchem et al., 2001), were identified through literature reviews of studies previously read. After the abstracts of the hits were screened, most were excluded for reasons such as being irrelevant to the current research topic, including other interventions/not including both self-management and

interdependent group contingency and being explorative rather than experimental research (see appendix 8.2). Twenty-two full-text papers were assessed for eligibility and seventeen excluded on similar grounds. Five remained to be reviewed. See appendices 8.1 and 8.2 for more information on the search strategy and reasons for exclusions.

2.5.4. **Assessment of study quality**

The five studies which met the inclusion criteria were critically evaluated using Gough's (2007) 'Weight of Evidence' (WoE) quality assessment tool (Figure 2.3). However, as this tool does not clearly define specific criteria with which to make the judgements, the author made judgements on the Weight of Evidence A based on criteria outlined by Trickey and Topping (2004), and Greenhalgh and Brown (2014). See appendix 8.4.

Weight of Evidence A	A generic judgement about the coherence and integrity of the evidence in its own terms.
Weight of Evidence B	A review specific judgement about the appropriateness of that form of evidence for answering the review question; the fitness for purpose of that form of evidence.
Weight of Evidence C	A review specific judgement about the relevance of the focus of the evidence for the review question in terms of type of sample, the type of evidence gathering or analysis, the context and ethical issues.
Weight of Evidence D	An overall judgement of quality and contribution towards answering the review question based on judgements of 'Weight of Evidence' A-C.

Figure 2.3: Descriptors for Gough's (2007) Weight of Evidence Framework.

Each study's overall (Weight of Evidence D) rating is outlined in Table 2.3. For more information on each study's ratings on the Weight of Evidence A-C and a breakdown of factors which led to the Weight of Evidence A rating, see appendix 8.4.

2.5.5. **Review of studies**

The following review will first examine research which measured the impact of self-management and interdependent group contingency combined and then examine the research which either examined them

separately or examined one and then measured the additive impact of the other.

Study	Participants	Intervention	Design	Key Findings	Weight of evidence
Chafouleas, Sanetti, Jaffery & Fallon (2012)	57 middle school pupils	Self-management THEN group contingency added	Multiple baseline with embedded changing criterion	On average, slight to moderate improvements in class-wide behaviour were observed.	MEDIUM
Davies & Witte (2000)	30 7-8 year-old pupils (n=4 with ADHD)	Self-management AND group contingency together	ABAB reversal design	Substantial decreases in inappropriate talking-out behaviour of ADHD pupils and their matched controls. Positive interdependence reported.	MEDIUM
Denune, Hawkins, Donovan, McCoy, Hall & Moeder (2015)	14 10-11 year-old pupils with EBD	Group contingency THEN self-management added	ABCBC withdrawal design	Study reports that group contingency was effective in reducing disruptive behaviour and increasing on-task behaviour, with no added benefit of including self-management however behaviour reached ceiling and floor levels following group contingency so the impact of self-management could not be measured.	HIGH
Glynn, Thomas & Shee (1973)	37 6-7 year-old pupils	Group contingency only THEN Self-management only	ABCACDEEAE design	All intervention conditions showed an increase in on-task behaviour compared to baseline conditions. Self-management showed slight improvement over group contingency.	LOW

Hoff & Ervin (2013)	64 6-7 year-old pupils	Group contingency THEN Self-management added	Multiple baseline across subjects design	Decrease in disruptive behaviour at the class-wide and individual level.	HIGH

Table 2.3: A summary of the studies selected for the systematic review.

2.5.5.1. Self-management and interdependent group contingency combined

Davies and Witte (2000), measured the impact of implementing self-management with interdependent group contingency on eight pupils ranging from 8-10 years-old (four with ADHD and four matched controls), within one classroom. An ABAB reversal design was used and the approaches were conducted for two months in a mainstream elementary school. The approaches aimed to reduce inappropriate verbalisations and measured this using event recording over 30-minute observations. The findings suggested that these approaches substantially decreased inappropriate verbalisations in all participants and increased positive interdependence.

The sample size of this study was small, making it difficult to reach strong conclusions and to generalise the findings to other populations. However the pupils were described in sufficient detail to support generalisability for other pupils with a similar profile (Horner et al., 2005). Also, four weeks into the study, three of the matched controls were moved into a different class and new matched controls were assigned. This may have affected the results, compromising the reliability of the findings, especially as the class size had reduced considerably. Inappropriate verbalisations may in fact have reduced because the teacher had greater capacity to attend to the target pupils. Additionally, as this study only measured one type of classroom behaviour, it is not possible to know what impact these approaches had on other behaviours. This study rated **medium** in Gough's (2007) Weight of Evidence quality assessment.

2.5.5.2. Impact of self-management or interdependent group contingency separately and combined

Glynn, Thomas and Shee (1973), measured the impact of interdependent group contingency on a class of thirty-seven 6-7 year-olds in New Zealand and then measured the impact of self-management, with interdependent group contingency removed. Data

was collected on the whole class as well as eight randomly chosen pupils, using a 10-second interval recording of whole-group and individuals' on-task behaviour. This study utilised a complicated ABCACDEAE design and the results suggested that on-task behaviour improved for individuals and the whole group compared to baseline levels, with self-management having a greater impact over interdependent group contingency.

These results must be interpreted with caution as the data collectors found it difficult to measure on-task behaviour, meaning the results may be inaccurate. In addition, parts of the intervention were not adhered to by the teacher and for some time, no data was collected due to the unavailability of observers. Furthermore, the sample size was limited as only one class participated in the research and generalisability is difficult as no information was reported about the demographics of the school population or the pupils included in the study. Furthermore, the view that self-management led to better outcomes in comparison to interdependent group contingency must be interpreted with caution as it is possible that interdependent group contingency which came before, may have led to lasting improvements which enabled self-management to have an additional impact which may not have been possible without group contingency's prior presence (ordering effects). This study rated **low** in Gough's (2007) Weight of Evidence quality assessment.

Denune, Hawkins, Donovan, McCoy, Hall & Moeder (2015) assessed the impact of implementing interdependent group contingency first and then adding self-management, on the engagement and disruptive behaviours of fourteen 12-15 year-old pupils with 'emotional and behavioural disorders' in an EBD school. Some pupils had diagnoses of ADHD, Oppositional Defiance Disorder and Post-traumatic Stress Disorder. The sample consisted of male and female participants who were White or Black. An ABCBC withdrawal design was used and the length of the interventions varied between six to nine weeks. Data was collected from more than one observer, using a 20-second momentary interval recording and a 20-second partial interval recording. The study reported decreases in disruptive behaviour and increases in on-task

behaviour following the implementation of interdependent group contingency, however the addition of self-management was not considered to have led to further improvements. Maintenance data collected a week later suggested that improvements in engagement continued, with slight increases in disruptive behaviour. Fidelity measures suggested that procedures were correctly followed by the teacher.

These results suggest that interdependent group contingency is effective for improving behaviour in class, however the conclusion that self-management has no impact may be inaccurate, as the levels of behaviour reached floor and ceiling levels following interdependent group contingency. As a result, the possible impact of self-management was undetectable. Furthermore, the maintenance data was only collected up to a week after the withdrawal of the interventions. As such, it is not possible to know what the long-term impacts of the approaches were. This study provided much information regarding the school and its participants, supporting the generalisability of the findings to similar populations but not necessarily mainstream pupils. Also, the sample size was small, limiting generalisability. This study rated **high** in Gough's (2007) Weight of Evidence quality assessment.

Chafouleas, Sanetti, Jaffery & Fallon (2012) measured the impact of implementing self-management first and then adding interdependent group contingency, on the preparedness, on- and off-task behaviour and homework completion of fifty-seven 12-13 year old pupils, across two mainstream classes. The school contained 20% of pupils from low income families and the participants of the study were White, Hispanic, Asian American or Biracial. The length of the interventions varied between four to six weeks and the study employed a multiple baseline design with a changing criterion element such that their target was gradually increased over time. Fidelity checks were conducted and on-going support offered to increase treatment integrity. Data was collected using a 15-second interval momentary time sampling procedure and partial interval recording. Pupils also filled out a Direct Behavior Rating form to measure their own behaviour and observers filled out a Systematic Direct Observation form to measure

engagement. These forms were adapted from the Behavioral Observation of Students in School (BOSS) schedule (Shapiro, 2013).

The findings from this study suggested that overall, there was a slight to moderate improvement on behaviour from implementing self-management only, to adding interdependent group contingency. However, behaviour was not rated prior to implementing self-management and so it is not possible to know the impact that self-management only, had on behaviour. The data from different classes highlighted varying degrees of improvement. Off-task behaviour was reported to have shown the greatest improvement, however the comparatively smaller improvement on on-task behaviour may be attributed to initially high levels of engagement at baseline, when only self-management was in place. Furthermore, the pupil and teacher ratings differed, with the teacher rating tending to report higher levels of engagement and greater improvement than the pupil rating, although the trends were similar. Generalisability of the findings is difficult due to a small sample size. This study rated **medium** in Gough's (2007) Weight of Evidence quality assessment.

Hoff and Ervin (2013) implemented interdependent group contingency first and then added self-management with 6-7 year olds, in four mainstream classrooms, however data from one class was excluded due to it not meeting a pre-specified criterion. As such, data was only collected in three classrooms which totalled sixty-four pupils. The researchers collected data on disruptive behaviours of the whole class and individual data on three 'at risk' pupils (one from each class, two with ADHD diagnoses). The study was conducted with a predominantly Caucasian population and few pupils from low income families. A multiple baseline design was employed and the interventions varied between eight to twenty-six sessions in length. Data on disruptive behaviour was collected using a 15-second partial interval recording.

The researchers reported improvements for the whole class as well as the three at-risk pupils, however they were unable to establish whether self-management or group contingency had the greatest impact. Data from Class 1 suggested that interdependent group contingency

decreased disruption compared to baseline and that self-management decreased this further, to floor levels. The degree of improvement was small however, as the baseline levels of disruption in this class were low. Data from Classes 2 and 3 suggested reductions in disruptive behaviour when interdependent group contingency was implemented, however self-management did not appear to reduce disruption further. These findings support Denune et al's (2015) view that self-management may not have an additive benefit to interdependent group contingency.

Unfortunately, this study did not report follow-up data in order to ascertain whether the improvements were maintained following the withdrawal of self-management. With a small sample size, generalisability of the findings is difficult. The integrity of the procedures was improved through fidelity checks and biweekly meetings with the researchers. Furthermore, this study to employed blind assessors, which reduced the likelihood of observer bias. This study rated **high** in Gough's (2007) Weight of Evidence quality assessment.

2.5.6. Approach procedures, data collection methods and data analysis

This review highlights the difficulty with building an evidence base for whole-class self-management and interdependent group contingency for classroom management. Only a limited number of studies have researched the area and all employ different procedures, data collection methods and data analysis procedures, making it difficult to effectively triangulate findings into a coherent conclusion. For instance, Davies and Witte (2000) required groups to move black dots onto group charts, as well as keep an individual record of their own behaviour for self-management, whereas pupils in Hoff and Ervin's (2013) study rated their behaviour on a 1-5 scale. In Denune et al's (2015) study, pupils simply wrote 'yes' or 'no' to whether they were following the rules at the moment they were stopped. In Chafouleas et al's (2012) study, pupils rated their behaviour for self-management only at the end of the lesson. For interdependent group contingency,

Chafouleas et al's (2012) study increased the criterion for receiving rewards, whereas Denune et al's (2015) study chose the criterion at random. No two studies implemented procedures in the same way. This makes the approaches flexible enough to be adapted to different age groups and classes, but makes it difficult to build an evidence base.

Furthermore, some studies observed individuals one at a time (Chafouleas et al., 2012; Denune et al., 2015; Hoff & Ervin, 2013), and others measured whole-class behaviour by taking auditory and visual sweep of the whole class (Glynn et al., 1973). The duration of intervals varied, as did the duration of the overall observation period. None of the reviewed studies used any standardised measure of behaviour or engagement which may have affected the reliability and validity of the results. With the exception of Davies and Witte's (2000) study, the findings tended to be analysed visually through percentage of overlapping data points (Chafouleas et al., 2012; Denune et al., 2015), or through calculating mean and variance (Glynn et al., 1973; Hoff & Ervin, 2013), which is argued to be unreliable (Ninci, Vannest, Willson, & Zhang, 2015), rather than finding statistical significance.

2.5.7. **Limitations of the review**

The limitations to this review are: firstly, due to publication bias, there is an increased likelihood that there are studies which found negative results but were not published and therefore were not included in this review (Dundar & Fleeman, 2014). Secondly, as Gough's (2007) Weight of Evidence tool does not give specific criteria with which to make judgements of quality, the author's judgements may have been too subjective. Additionally, the researcher may have excluded studies that somebody else may have included. Finally, it was not possible to do inter-rater checks, which would have reduced the possibility of bias in quality assessment (Greenhalgh & Brown, 2014).

2.5.8. **Summary of the review**

The studies reviewed ranged from low to high quality and none were conducted in the UK, making generalisation to the UK population

difficult. In all studies, the class or pupils served as their own controls and none used a group design to control for maturation.

Even though all of the studies reported efficacy with using whole-class self-management and interdependent group contingency for reducing disruptive behaviour and increasing engagement, maintenance checks were not conducted to identify whether the improvements were maintained over time. The use of small sample sizes, varied age groups and different procedures impacts on the ability to build an evidence base for these approaches due to difficulties in directly comparing the results. The current evidence suggests that these approaches are promising for classes in different phases of school but more research of better quality is needed, particularly in the UK. Future studies must replicate the same procedures with the same population demographics and in similar age groups, as well as reducing the risks posed by floor and ceiling, and ordering effects. The use of group designs, standardised measures and statistical analyses would also be welcomed.

This review highlighted that two questions remain to be answered: firstly, what is the impact of whole-class self-management on classroom behaviour, compared to only using interdependent group contingency? Secondly, does adding self-management to interdependent group contingency, or adding interdependent group contingency to self-management lead to additional improvements to behaviour than either do alone? Furthermore, owing to the increased concerns around behaviour and exclusions in primary school (Department for Education, 2014), research which focuses these approaches in the primary age phase, would be welcomed.

2.6. **Research questions**

The present study aims to answer the following research question:

How effective is a) a whole-class self-management and b) an interdependent group contingency approach, in terms of improving overall behaviour in junior school classes?

This research question will be addressed through investigating the following sub-questions:

1) *What impact does whole-class self-management have on off-task and disruptive behaviours, as well as on the general behaviour of the whole class, as compared to control groups?*
2) *What impact does interdependent group contingency have on off-task and disruptive behaviours, as well as on the general behaviour of the whole class, as compared to control groups?*
3) *Which approach (self-management or interdependent group contingency) is most effective in reducing off-task and disruptive behaviours, and in improving general behaviour?*
4) *Is there an added benefit to combining self-management with interdependent group contingency, with regard to off-task and disruptive behaviours, as well as general behaviour?*

3. Methodology

The previous chapter outlined the key literature around the topic of behaviour and classroom management, and identified four research sub-questions to be investigated.

The current chapter will begin by discussing the importance of, and difficulties surrounding, real world research and evidence-based practice. Next, key philosophical assumptions, theoretical paradigms and research designs which guide social research will be discussed before the rationale for the chosen paradigm and design of this study is outlined. The characteristics of the participants and school in this study will be highlighted. The methods employed in this study will be outlined, including how the sample was recruited, how the teachers and pupils were trained and how the approaches were implemented. The measures used for collecting data will be described and finally, issues of data quality and ethical considerations will be discussed.

3.1. Real world research

The current study is an example of real world research. Real world research "focuses on problems and issues of direct relevance to people's lives, to help find ways of dealing with the problem or of better understanding the issue" (Robson, 2011, p. 4) and is conducted in a variety of settings within communities (Gray, 2014). Conducting research of this kind is valuable in education and psychology, as the findings generated by research conducted in laboratories have often not been replicated in real world settings (Robson, 2011).

It can be difficult however, to gain access to settings due to the differing agendas of researchers, stakeholders and gatekeepers (Gray, 2014). Additionally, real world research is often constrained by the time and funds available (Robson, 2011). Often, an inter-disciplinary approach which draws upon areas such as psychology, sociology and philosophy is needed (Gray, 2014), and so the researcher must be knowledgeable and skilled in a variety of research methodologies (Robson, 2011). The greatest threats to real world research however,

are the threats to internal validity and reliability due to the difficulty of controlling for extraneous variables (Mertens, 2005; Robson, 2011).

3.2. **Evidence-based practice**

Issues of variability in service delivery has been of concern in Educational Psychology (Department for Education and Employment, 1998). Evidence-based practice aims to ensure that "clinical decisions [are] based on the best possible evidence of effectiveness" (Department of Health, 1998, p. 2). As such, there has been increased interest in Educational Psychologists (EPs) applying psychology in an evidence-based way by ensuring that research evidence is used alongside professional judgement when suggesting interventions (Frederickson, 2002).

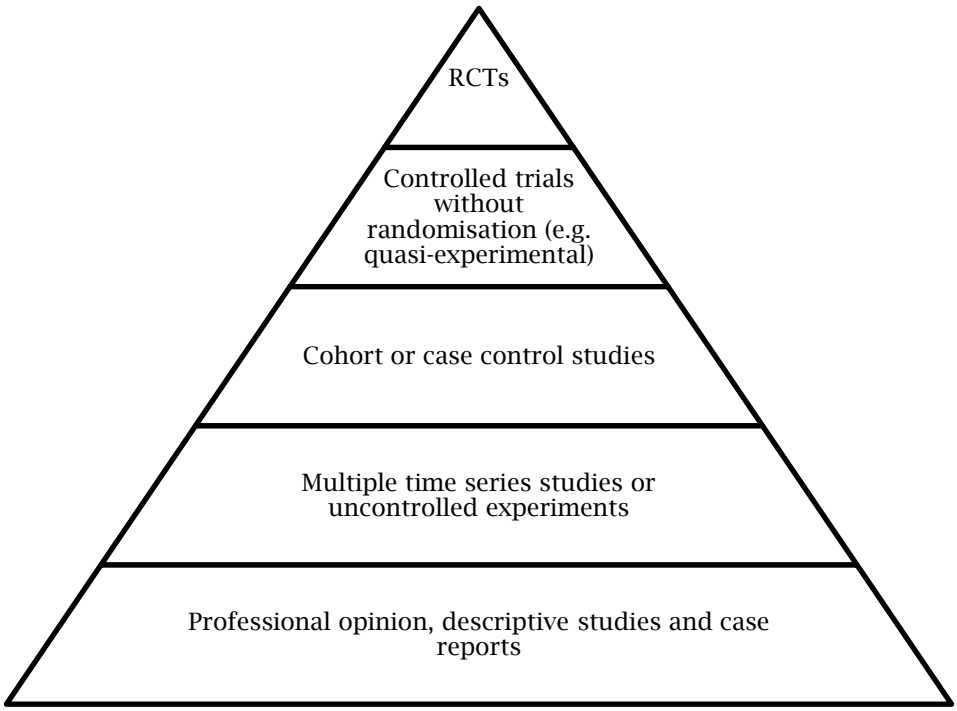


Figure 3.1: An example of a hierarchy of evidence from health, adapted from Concato, Shah and Horwitz (2000).

The traditional hierarchy of evidence (Concato et al., 2000), developed for use in health (N. B. Anderson, 2006; Frederickson, 2002) (Figure 3.1), views randomised control trials (RCTs) as the 'gold standard' form of evidence (Robson, 2011), however Roth and Fonagy (2005) argue that a tension exists between research with high internal validity which

is tightly controlled (as with RCTs), and its relevance and generalisability in real world settings. Quasi-experimental designs have been considered a second-best choice to RCTs (Frederickson, 2002; Robson, 2011), and are often more appropriate in the real world context (Robson, 2011). The current research aims to add to the evidence-base of whole-class approaches to classroom management, with a view to informing the work of, and recommendations made by EPs.

3.3. **Philosophical paradigms in research**

A paradigm is "the ways that we think about research and the world" (Thomas, 2013, p. 105), and "how we seek knowledge and how we use it" (Thomas, 2013, p. 106). Guba and Lincoln (1994) outlined three questions that can be asked in order to identify the basic beliefs held by the researcher and therefore their philosophical standpoint when designing research:

The ontological question: *"What is the form and nature of reality and, therefore, what is there that can be known about it?"* (p. 108).

The epistemological question: *"What is the nature of the relationship between the knower or would-be knower and what can be known?"* (p. 108).

The methodological question: ***"****How can the inquirer (would-be knower) go about finding out whatever he or she believes can be known?"* (p. 108).

Below, three major theoretical paradigms are outlined and discussed in relation to these questions.

3.3.1. **Positivism and post-positivism**

Positivism has been a popular paradigm since the mid-nineteenth century (Crotty, 1998; Thomas, 2013). This paradigm assumes that there is one reality that can be known (ontology), that the researcher is separate to the object under investigation and can study it objectively

without mutual influence (epistemology), and that quantitative and carefully controlled experimental methods driven by hypotheses should be used to do so (Gray, 2014; Guba & Lincoln, 1994; Mertens, 2005). Positivists saw the natural and social worlds as both being governed by scientific laws, which could be investigated empirically (Gray, 2014). Positivism however has been widely criticised for making excessive claims around its level of objectivity (L. Cohen, Manion, & Morrison, 2000) and conviction in the robustness of the results it obtains (Crotty, 1998).

In response to these criticisms, post-positivism emerged (Gray, 2014). A post-positivist stance takes the view that a reality exists, however our ability to comprehend it is limited by human limitations (ontology) (Guba & Lincoln, 1994; Mertens, 2005; Robson, 2011). This paradigm rejects the belief that the researcher and object of study do not mutually influence each other, however maintains that objectivity is important and so values and biases should be prevented from influencing the research (epistemology) (Guba & Lincoln, 1994; Mertens, 2005). Post-positivists also recognised that rigorous scientific methods were not appropriate in the social sciences and so adapted them to be applicable in real world contexts, such as through quasi-experimental designs, where participants are not randomly allocated to conditions (Mertens, 2005).

3.3.2. **Constructivism**

Critics have argued that positivist research inaccurately simplifies human behaviour to being passive and controlled rather than influenced by intention and freedom, and that humans create theories about the world and themselves, which influence their behaviour (L. Cohen et al., 2000). In line with this, the constructivist paradigm views reality as socially constructed and therefore it is argued that multiple constructions can be acquired, which may differ to others' constructions (ontology) (Gray, 2014; Mertens, 2005). The idea that there is one reality, is rejected (Mertens, 2005). This paradigm views the researcher and the object of the research as being linked into an interactive, mutually influential process where the researcher's values

are made explicit, and the outcomes and interpretations of the research are jointly constructed and context dependent (epistemology) (Guba & Lincoln, 1994; Mertens, 2005). In accordance with the philosophical stance of this paradigm, qualitative data are collected through methods such as interviews and observations (Mertens, 2005).

3.3.3. **Pragmatism**

The pragmatic paradigm rejects the debates around a positivist vs. constructivist ontology and an objective vs. subjective epistemology (Coe, 2012). Rather than searching for a 'true reality', pragmatists argue that 'effectiveness' is more important; namely, determining what is useful to the research (ontology) (Lodico, Spaulding, & Voegtle, 2010). The relationship between the researcher and object of research depends upon what is deemed appropriate in order to answer the research question (epistemology) (Lodico et al., 2010; Mertens, 2005). As such, the chosen methodology can be qualitative and/or quantitative, depending on the needs and purposes of the research (Creswell, 2013; Mertens, 2005; Teddlie & Tashakkori, 2009).

Despite on-going debates around which philosophical stances should be taken in research, it is argued that none of the paradigms can be proven as being correct (Guba & Lincoln, 1994). As all paradigms are human constructions, it is argued that they are all liable to human error (Guba & Lincoln, 1994).

3.4. **Research designs**

3.4.1. **Fixed designs**

Robson (2011) distinguishes between fixed and flexible research designs in the social sciences. Fixed designs align themselves with positivist and post-positivist paradigms (Robson, 2011), and are theory-driven. They require the variables to be examined, as well as the design, procedures and choice of data analysis to be stated at the outset (Robson, 2011). The findings from these studies inform theories and/or models (Robson, 2011). It has been argued however, that fixed design research is unable to detect the subtleties and complexities of

human behaviour at the individual level (Dixon, Singleton Jr, & Straits, 2016; Robson, 2011), which is problematic when trying to understand human behaviour in the real world. Furthermore, fixed design researchers are 'locked' into their chosen design and lack the flexibility to adapt in light of unexpected findings (Dixon et al., 2016). Examples of fixed designs include true experiments, single-case experimental designs and quasi-experimental designs (Creswell, 2014; Robson, 2011). A quasi-experimental design was chosen for this research, given that experimental designs are considered to be the most robust type of evaluation studies for evaluating approaches or interventions (Robson, 2011). As such, they are discussed in more detail below.

3.4.1.1. Quasi-experimental designs

Experimental designs aim to establish a cause and effect relationship by deliberately manipulating one variable (the independent variable, IV) and measuring its impact on another variable (the dependent variable, DV) using pre- and post-measures, in comparison to a control group who receives a different or no intervention (L. Cohen et al., 2000; Thomas, 2013). As previously discussed (see Section 3.2), although randomly allocating participants to conditions in RCTs is considered the best option in that it allows key extraneous variables to be controlled for, strengthening the causal link (Robson, 2011), this is often not feasible in the real world. As such, predetermined groups are often allocated to conditions, leading to non-equivalent groups which may increase the risk of regression towards the mean on post-test measures (L. Cohen et al., 2000), and a lack of control over key variables (Mertens, 2005; Robson, 2011; Thomas, 2013) which increases the threats to internal validity (the risk that the effects measured are due to factors other than the independent variable) (Robson, 2011; Thomas, 2013). However, some have reasoned that controlling for a large number of variables can lead to an oversimplification of the phenomenon, making it difficult to understand how it operates in the real world (Mertens, 2005).

3.4.2. **Flexible designs**

Unlike fixed designs, flexible designs allow the design to develop and change during the course of the study (Robson, 2011). Researchers are likely to construct a design which fits their research question, rather than choose a specific, pre-existing design (Robson, 2011). This research tends to collect a variety of qualitative data, such as interview data (although some quantitative data may be included) and these present multiple realities which focus upon participants' subjective views rather than trying to obtain a 'true' answer (Robson, 2011). The research begins with a particular idea or problem to be understood (Robson, 2011), but the aim is not to identify causal relationships or to compare groups (Robson, 2011). This design aligns itself with the constructivist paradigm (Robson, 2011). Some argue that flexible designs may be more valid in real world research as it is possible to obtain a holistic, in-depth view of a situation (Dixon et al., 2016), however it is then difficult to generalise the findings from the individual, to wider groups (Dixon et al., 2016).

3.4.3. **Mixed method designs**

In mixed method research, qualitative and quantitative characteristics are both present in the design, data collection and analysis (Teddlie & Tashakkori, 2009). These designs align themselves with the pragmatic paradigm (Mertens, 2005; Teddlie & Tashakkori, 2009) and include sequential explanatory, sequential exploratory and sequential transformative designs (Robson, 2011). The view is to triangulate the data from a variety of sources and to analyse it in different ways to gain multiple perspectives (Mertens, 2005). A difficulty with mixed method research is that the researcher must be fluent with both fixed and flexible research designs (Mertens, 2005).

3.5. **The current research**

3.5.1. **The paradigm**

The current study sought to evaluate the efficacy of two whole-class approaches to classroom management in reducing disruptive

behaviour and increasing engagement. As such, this research aimed to establish a cause and effect relationship and to arrive at a 'true' answer about the effectiveness of the approaches, which could be generalised to the wider population and which would contribute towards building an evidence-base. These aims are consistent with a post-positivist standpoint, which aims to maintain objectivity through adapting the scientific method for use in real world settings whilst recognising potential biases in the findings (Robson, 2011). As such, a post-positivism paradigm was chosen.

3.5.2. **The design of the current Study**

Consistent with a post-positivist standpoint, an experimental design was chosen for this research, as such designs are considered most appropriate for evaluation studies (Robson, 2011). Given that pupils in this study were already placed in predetermined classes, it was not possible to randomly allocate participants to conditions. As such, a quasi-experimental design was used, which ranks highly on the hierarchy of evidence (Concato et al., 2000), and arguably, has better ecological validity than RCTs (Robson, 2011).

This study employed two phases. Phase A investigated the impact of 'self-management alone' and 'interdependent group contingency alone' on engagement and disruptive behaviours of the whole class, seeking to determine which was most effective alone. Phase B investigated whether combining the approaches led to further reductions in disengagement and disruption, in order to determine whether it is best to employ both approaches *in combination* (see Table 3.1). Data was collected at three time points, where Time 1 and Time 2 measures served as pre- and post-measures for Phase A, and Time 2 and Time 3 measures served as pre- and post-measures for Phase B. In order to control for ordering effects (the idea that being exposed to one treatment before another could impact the results differently than if they had been administered in reverse order) (Mertens, 2005), the research was designed so that one class would receive self-management before adding interdependent group contingency, and another class would receive interdependent group contingency before

adding self-management. Unfortunately, due to long-term staff absence and subsequent teacher resignation, Time 3 data could not be collected from one of the four classes that participated (Class 2) and therefore could not be included in the research. This meant that research question 4 could only be answered by examining the impact of adding interdependent group contingency to self-management; not the impact of adding self-management to interdependent group contingency.

The approaches in Phase A and Phase B were both implemented for four school weeks before Time 2 and Time 3 data were collected, respectively. The Phase A approaches (self-management and interdependent group contingency) continued to be run during the two weeks of Time 2 data collection, which meant that by the end of the Time 2 data collection, the approach had been running for six school weeks before Phase B began. Similarly, the Phase B approach (self-management with interdependent group contingency) continued during the two weeks of Time 3 data collection, meaning it was implemented for six weeks in total, in Class 1. For Class 1, self-management from Phase A continued into Phase B. For Class 2, interdependent group contingency from Phase A continued into Phase B. Previous research has noted improvements after implementing the approaches for as little as two school days (Davies & Witte, 2000; Denune et al., 2015; Glynn et al., 1973), although the majority implemented the approaches for around four weeks (Chafouleas et al., 2012; Davies & Witte, 2000; Glynn et al., 1973; Hoff & Ervin, 2013). As such, four weeks was considered an appropriate length for this research, to keep within manageable timescales. Unfortunately, due to staff illness prior to the Christmas holidays, Class 2's Time 2 data was delayed by five weeks. As such, Class 2's Phase B commenced when the other classes were three school weeks into Phase B.

Table 3.1 outlines the design of the research. The choice of measures is discussed in detail in Section 3.5.8.

Class	Time 1 Measures (2 school weeks)	Phase A (4 school weeks)	Time 2 Measures (2 school weeks)	Phase B (4 school weeks)	Time 3 Measures (2 school weeks)
1 (Experimental)	Teacher-completed SDQ 3x classroom observations	Self-management (SM)	Teacher-completed SDQ 3x classroom observations Phase A experimental approaches (SM, IGC) and rule reminder continued to be run at this time	Self-management with Interdependent group contingency	Teacher-completed SDQ 3x classroom observations Phase B approaches (SM with IGC) continued to be run at this time
2 (Experimental)		Interdependent group contingency (IGC)		Interdependent group contingency with Self-management	---
3 (Waitlist Control)		Rule reminder		Rule reminder	Teacher-completed SDQ 3x classroom observations Rule reminder continued to be run at this time
4 (Waitlist Control)		No change		No change	

Table 3.1: A table to illustrate the design of the current study.

3.5.3. Stakeholders

3.5.3.1. University of Nottingham

The completion of this research project formed part of the course requirements for the Doctorate in Applied Educational Psychology at the University of Nottingham.

3.5.3.1. The researcher

The researcher was a stakeholder in this project as it was a necessary course requirement for the completion of the three year Doctorate in Applied Educational Psychology.

3.5.3.2. The Local Authority

At the time of this project, the researcher was working as a Trainee EP at the Educational Psychology Service in a large county authority. This research project was discussed and agreed with the Principle EP, who considered this area to be of potential benefit to the Service. As such, it was agreed that the findings of this research would be shared with the EPs of the Service and possibly be included within the menu of training offered to schools, if found to be efficacious.

3.5.3.3. School staff, parents and pupils

Other stakeholders included: the head teachers of the school, the teachers who participated in filling out questionnaires, attending training and implementing the approaches in their classes, the parents who consented to individual data being collected on their child and the pupils who were observed and were the subjects of the questionnaire data.

3.5.3.3.1. Engagement

The researcher sought three-form entry primary schools in order to have three age-equivalent groups for the research, as age-equivalent cohorts are considered most comparable (Shadish, Cook, & Campbell, 2002). Stakeholder engagement letters (appendix 8.6) were sent to

three-form entry primary schools and followed up with telephone calls, with no success. The participating school was recruited following liaison between the head teacher and their previous link EP (a colleague of the researcher), and after the researcher and head teacher had met to discuss the research further.

During the meeting with the head teacher, the researcher explained the research aims as per appendix 8.7. The researcher explained what would happen in the target classrooms if the school participated, what the time and other commitments would be for the teachers, the length of the entire project (see appendix 8.8), the benefits of participating and the commitments and limits of confidentiality and anonymity on the part of the researcher (see appendix 8.7).

During this meeting, the researcher explained that she wished to work with all teachers and children within one of the year groups in the junior school, where many of the children displayed off-task and disruptive behaviour in a particular lesson (preferably literacy or numeracy, which would allow the approach to be implemented daily, in the same lesson). It was imperative that the children had not received these approaches or something similar before.

The head teacher was informed that the senior leadership team and teachers in collaboration could select which cohort they felt would benefit from these approaches. However, it was made clear that with teacher consent (see appendix 8.8), observations would be conducted in those classes to check that baseline levels of off-task and disruptive behaviour were sufficiently high for the approaches' efficacy to be detectable. If they were not, those classes could not be used. The researcher explained that it would be necessary for her to meet with the teachers to explain the research aims and to answer any of their questions before seeking informed consent for their participation (see appendix 8.8). It was also stated that the teachers must participate of their own free will and must not feel coerced. Furthermore, they would have the right to withdraw at any point, with no consequences from the school or elsewhere. The head teacher was keen to participate and she signed the head teacher consent form (appendix 8.7).

Before the end of term, the head teacher informed the researcher that Year 4 had been selected and that the teachers were willing to meet with the researcher in September 2016, in order to learn more about the research and decide whether they wished to participate.

The teachers met with the researcher in September 2016 and gave informed consent for their participation (appendix 8.8). As this school was four-form entry it allowed for a second control condition (rule reminder) to be included, which enabled the researcher to investigate whether any changes in the experimental groups may simply be due to daily rule reminders, rather than specifically the elements of self-management or interdependent group contingency.

3.5.4. School and community context

This research was conducted in a four-form entry junior school, situated in the outskirts of a town in the Midlands, which is ranked at the 47th percentile on the Index of Multiple Deprivation (UK Local Area, 2016). The school had over 450 pupils on roll, of whom 18.5% had special educational needs, 13.8% received free school meals, 14% had English as an additional language and 83% were White British in comparison to 17% who were of a minority ethnic mix. 53.5% of pupils were male compared to 46.5% females.

3.5.5. Sampling procedure

This study recruited an opportunity sample. In consultation, the head teacher indicated that Year 4 were of most concern and felt that they would benefit from the classroom management approaches. 30-minute observations were conducted in each class to establish whether baseline levels of disengagement and disruption were sufficient for improvement to be detectable. These initial checks highlighted sufficient levels of disengagement and disruption. The teachers reported that approaches of these kinds had not previously been used with the pupils. As such, these classes were deemed suitable.

3.5.6. **Participants**

3.5.6.1. Classes

Table 3.2 provides contextual information about the classes that participated in this study.

	Class 1	**Class 2**	**Class 3**	**Class 4**
Teacher gender	Male	Female	Female	Male
Number of years teaching experience	12	12	4	8
Number of pupils included	30	29	27	26
Number of pupils excluded from SDQ data	0	1 female (Pupil with no SEN – opted out by parents)	1 female (Pupil with moderate learning difficulties – not in class during target lessons)	1 male (Pupil with autism – not in class during target lessons)
Pupil gender	Male: 17 Female: 13	Male: 18 Female: 12	Male: 18 Female: 10	Male: 20 Female: 9
Pupil age range	9 years 1 month – 8 years 2 months	9 years 1 month – 8 years 3 months	9 years 0 months – 8 years 1 month	9 years 1 month – 8 years 1 month
Average pupil age	8 years 4 months	8 years 8 months	8 years 6 months	8 years 8 months

	Class 1	Class 2	Class 3	Class 4
Pupils with SEN	0	0	2 Moderate Learning Difficulties (n=2)	4 Autism (n=1) Moderate Learning Difficulties (n=3)
Pupils with English as an Additional Language	10	5	6	5
Pupils Looked After	0	0	0	0
Pupils on free schools meals	5	6	1	4

Table 3.2: A table outlining the contextual information of the classes included in this study.

3.5.7. Implementing the approaches

3.5.7.1. The procedures

Given the varied procedures for implementing self-management and interdependent group contingency, the researcher adapted those used by Hoff and Ervin (2013), and Davies and Witte (2000), as their procedures were conducted on a mainstream, primary-aged sample. Furthermore, the researcher ensured that the procedures met the definitions of what constitutes 'self-management' and 'interdependent group contingency', such that at least two of the following were present: self-monitoring, self-evaluation and positive reinforcement (Mitchem & Young, 2001), a rule was specified, a criterion was outlined for the group to meet, the reward was agreed and was administered when the criterion was met (McKissick et al., 2010; C. H. Skinner, 2004).

The procedures used in this study are described briefly in Table 3.3 and in more detail in appendix 8.10.

Conditions	Details
Self-management only	**Pupil training:** Teacher trains the pupils on how to self-rate accurately. Classroom rules are outlined and examples of rule-breaking and rule-following behaviours are discussed. **Procedure:** The lesson starts with a rule-reminder. Every 15 minutes, pupils are stopped by the teacher and asked to rate, on a scale of 0-4, how well they were following a class rule chosen at random: 0=not at all; 1=some of the time; 2=half of the time; 3=most of the time; 4=all of the time. At the end of the lesson, the pupils rate the whole class on a randomly chosen rule. The majority vote is compared to the teacher's rating of the whole class. Pupils share their ratings with a partner and give feedback to each other.
Interdependent group contingency only	**Pupil training:** The class rules are outlined and discussed. The contingency for earning group rewards is shared. Ground rules are set to inform pupils to be supportive and respectful of each other. **Procedure:** The lesson starts with a rule-reminder. At the end of the lesson, the teacher chooses a rule at random and rates the class, on the 0-4 scale, on how well they followed the rule. The points are marked on a graph and will lead to a previously agreed reward once the pre-determined criterion is reached.
Interdependent group contingency & self-management	**Pupil training:** As per the procedures above. **Procedure:** (As per the procedures above) The lesson starts with a rule-reminder. The children rate their behaviour every 15 minutes on a randomly chosen rule. At the end of the lesson, the pupils then rate the whole class on a randomly chosen rule and the majority vote is compared to the teacher's whole-class rating. Where the teacher's and pupils' rating match, the pupils are awarded those points, plus a bonus point for matching the teacher's score. Where there is a difference of one point between the pupils and teacher, the class are awarded the teacher's points. If the difference is more than one, no points are awarded. The points are marked on a graph and will eventually lead to the pre-negotiated reward.

Conditions	Details
(Rule Reminder) Waitlist Control	The teacher trains the children on rule following and rule breaking behaviour only. The teacher begins the lesson every day by outlining/reminding the pupils of the classroom rules.
(No Change) Waitlist Control	The teacher conducts the lessons as usual. No training on the rules given.

Table 3.3: A table to outline the procedures for the classroom approaches used in this study.

The rules were chosen at random to reduce the risk that pupils would only follow the rule they knew would earn them points (McKissick et al., 2010). Peer feedback was also implemented in order to give the pupils an additional opportunity to reflect on their own behaviours and improve the accuracy of their ratings (Mitchem & Young, 2001; Mitchem et al., 2001). The rules were consistent across all three classrooms and were developed by the teachers during the initial training session. These were:

- *Be respectful to the teacher and to other children*
- *Stop and look at the teacher when the teacher is talking*
- *Follow teacher instructions straight away*
- *Allow other children to learn*

Unfortunately, due to the time limitations and capacity issues, it was not possible to pilot the approaches prior to commencing the research. Similar approaches have, however, been implemented successfully in other studies (Chafouleas et al., 2012; Davies & Witte, 2000; Denune et al., 2015; Hoff & Ervin, 2013), decreasing the likelihood that there would have been significant pragmatic difficulties. Furthermore the teachers felt that the procedures would be understood by the pupils.

3.5.7.2. Implementation

It was agreed with teachers that the approaches would be conducted daily, during one afternoon lesson, as they felt that the pupils were less engaged after lunch. The decision to implement the approaches in only one lesson per day was made because interdependent group contingency is argued to be most effective when administered for short periods such as in one lesson (Maggin et al., 2012). Also, studies included in the systematic review reported success when self-management and interdependent group contingency were implemented in only one lesson per day (Denune et al., 2015; Glynn et al., 1973; Hoff & Ervin, 2013). Furthermore, research on self-management has suggested that using the approach in only one lesson each day can lead to generalised improvements in other, non-targeted sessions (Amato-Zech, Hoff, & Doepke, 2006; Wolfe, Heron, & Goddard, 2000). It was also felt that these approaches may be more acceptable to teachers and lead to increased fidelity if they needed only to be implemented within a single lesson, daily.

The 13.30-14.30 lesson was originally targeted in order to control for time-of-day variables, however, it was not always possible to implement the approaches at that time. For instance it was impractical for pupils to carry a self-rating sheet in P.E. On those occasions, the approaches were implemented during the 14.30–15.30 lesson. As the pupils were in mixed ability classes for only the afternoon lessons, these groupings were considered to be potentially most equivalent. An additional advantage to targeting the afternoon lessons was that it meant the teacher who knew the pupil best (their form tutor) would complete the questionnaire data on their general behaviour, increasing the reliability of the data.

3.5.7.3. Teacher and pupil training

The teachers received training first for Phase A and then at a later date, for Phase B. Phase A training was delivered by the researcher to the teachers in Classes 1, 2 and 3 via Microsoft PowerPoint© presentation (see appendix 8.11). Each teacher received only the input relevant for

their approach in order to reduce the risk of diffusion of treatment. The first training session lasted around 90 minutes. Due to staff illness, only the teacher in Class 1 (self-management) received the training for Phase B in December 2016. The teacher in Class 2 (interdependent group contingency) received the training in January 2017. This training was delivered via Microsoft PowerPoint© presentation and lasted around 60 minutes (see appendix 8.12). During both training sessions, the teachers were able to ask questions and the first lesson in which the new approach was implemented was observed by the researcher or a psychology assistant (PA), to ensure it was appropriately delivered. The researcher supplied the rating sheets and the goal sheet (see appendix 8.13). The teachers used their own timers.

3.5.7.4. Fidelity checks

Four PAs in the Educational Psychology Service supported the researcher in conducting fidelity checks. They received in total, three hours of training from the researcher on the approaches and how to conduct these checks (see appendix 8.15 for the fidelity checklists used). No diffusion of treatment was observed in the waitlist control classes and the approaches were conducted mostly to fidelity in the experimental classes. Occasionally, a step or two was forgotten. Fidelity was also checked informally whilst conducting observations. Observations in the self-management class suggested that rule reminders were not consistently being given at the start of the lesson, however the pupils did appear confident with the approach and were sharing their ratings with a partner. See appendix 8.16 for more information on the outcomes of the fidelity checks conducted.

3.5.7.5. Timeline for the research

All elements of this research project were conducted between March 2016 and March 2017. See appendix 8.20 for a detailed timeline of the research process.

3.5.8. **Measures**

This study sought to assess the impact of self-management and interdependent group contingency on levels of off-task (disengagement) and disruptive behaviour of the whole class, both within the target lessons and more generally in school. Changes within the target lessons were measured through structured classroom observations. General changes in behaviour were measured through the Strengths and Difficulties Questionnaire (SDQ) (Goodman, 1997).

3.5.8.1. Strengths and Difficulties Questionnaire

Teachers were asked to fill out the SDQ for each pupil as pre-test and post-test measures in Phases A and B of the study (appendix 8.17). The questionnaire is a brief behavioural screening tool measuring emotional symptoms, behavioural problems, hyperactivity and concentration problems, peer problems and prosocial behaviour. For the purposes of this research, only the 'behavioural problems' and 'hyperactivity and concentration problems' subscale scores were analysed as the others did not relate to the research questions. The SDQ was chosen as a measure of generalised behaviour and self-regulation as it is a widely used and respected tool. It is also easy and quick to administer, taking only five minutes per pupil. Furthermore, it scores highly for internal consistency, test-retest stability (Goodman, 2001; Stone, Otten, Engels, Vermulst, & Janssens, 2010) and has reportedly strong psychometric properties (Stone et al., 2010). It has also been found to be a more accurate measure for detecting inattention and hyperactivity than the Child Behaviour Checklist (CBCL) (Goodman & Scott, 1999). As per the consent procedures, teachers applied a code to each questionnaire instead of a name, to ensure anonymity and the teachers were given two school weeks at each time point to complete the questionnaires.

A potential limitation of using the SDQ is that it is a subjective measure completed by the teacher, which could be prone to observer error (Robson, 2011). The advantage of obtaining teacher scores is that

they were able to give a general overview of the pupil's behaviour across lessons, which the researcher could not have directly obtained.

3.5.8.2. Observation schedules

A 'gold standard' classroom observation schedule for categorising disengagement and disruption does not exist (Briesch, Hemphill, Volpe, & Daniels, 2015). Previous research has tended to use sampling procedures which involve observing for intervals of a few seconds and categorising one behaviour per interval (Clemens, Shapiro, & Seibert, 2013). Examples include a partial interval recording procedure (behaviour can be present at any time during the interval, for it to be classed as present), and a momentary time sampling procedure (observer notes only the behaviour that was occurring at the end of the interval) (Briesch et al., 2013; Chafouleas et al., 2012; Clemens et al., 2013; Denune et al., 2015; Hoff & Ervin, 2013; Yoder & Symons, 2010), to observe individuals or groups in a random or fixed order (Briesch, Hemphill, et al., 2015). Sampling procedures make data collection simpler and more efficient (Clemens et al., 2013), however, they provide only estimates of the actual occurrence of behaviours and therefore may be inaccurate (Clemens et al., 2013; Rapp, Colby-Dirksen, Michalski, Carroll, & Lindenberg, 2008).

A partial interval recording was chosen for this study, as it is well-suited to observing behaviours of low, moderate and high frequency, which vary in duration and research suggests that it more accurately documents behaviour frequency than momentary time sampling (Clemens et al., 2013; Rapp et al., 2008; Yoder & Symons, 2010). In order to choose the most reliable and valid approach, the observation schedule was piloted.

3.5.8.2.1. Piloting for inter-observer agreement

A number of sampling procedures were piloted with an EP colleague, in a different school, to measure IOA, including a 15-second partial interval recording to observe groups of pupils and individuals in a fixed order according to seating plans. The IOA for observing groups was very low as it was difficult to observe so many individuals at once,

whereas observing individuals was much more feasible and resulted in 96.1% IOA with a Cohen's kappa coefficient of 0.917 (almost perfect agreement) (Viera & Garrett, 2005). It was not possible to measure IOA for disruptive behaviours as the class did not present much disruption; only one instance of disruption was observed, although this was observed by both observers.

3.5.8.2.2. Categorising disruptive and on- and off-task behaviour

Prior to piloting the observation schedule, on- and off-task behaviours (measuring engagement), and disruptive behaviours were operationalised according to prior research (Chafouleas et al., 2012; Denune et al., 2015; Ennis et al., 2016; Hoff & Ervin, 2013; T. Johnson et al., 1996; McKissick et al., 2010; Thorne & Kamps, 2008). Through piloting, these codes were adapted, in consultation with the EP colleague, to ensure categories were clear. This led to the following operationalised behaviour categories used within this study:

A child was considered **on-task** if they were yawning, glanced away briefly, were looking at the teacher/board/other pupils who were speaking about the task, answering questions, writing when asked to by the teacher, talking/calling out about the work, writing ideas down and waiting for teacher attention.

A child was considered **off-task** if they looked away from the teacher/work for more than 3 seconds (or looked away several times briefly), talked/shouted out while the teacher was talking, were walking around the classroom (exc. getting a resource), were fidgeting with objects while not looking at the teacher/work, were doodling, were not doing the set task, or were disrupting others while waiting for teacher attention.

Through discussion with the EP colleague, it was decided that **disruptive** behaviour would be a subcategory of off-task behaviour, in which the off-task behaviour disrupts or interrupts the learning of others by causing them to stop or look away from their work. This

definition is consistent with the ones used by Hoff and Ervin (2013) and Denune et al., (2015).

It was difficult to decide whether pupils talking was on- or off-task, when it was not possible to hear what was being said. Through discussion with the EP colleague, it was decided that the best option would be to base the decision on visual clues such as laughing and looking elsewhere. As such, it was collaboratively decided that on-task talking would be defined as talking where the pupil looked at their work/resources while talking, or was pointing to their work/resources while talking. Off-task talking was defined as talking where there was also laughter, or where the pupil was pointing/looking somewhere other than at their work, or where the pupils talking looked only at each other without referring back to their work.

3.5.8.2.3. Observation procedures

Following piloting of the observation schedule (see appendix 8.18), it was decided that 30-minute observations would be conducted using a 15-second partial interval recording to observe the pupils in a fixed order based on a seating plan. When the pupils sat on the carpet, they were observed in order from the researcher's left field of vision, across the room to the right. At the end of each interval, five seconds was used to record the behaviour category for that interval. If an off-task interval was considered to be disruptive, this was tallied on the observation sheet in order to keep a frequency count of disruptive intervals. Later, this was noted within the interval box to aid IOA checks.

At each data collection time point, the classes were observed on three separate days, across a two-week period. This was done in order to obtain a more reliable measure of the class' overall engagement, by obtaining a mean score across learning activities/days (Yoder & Symons, 2010). The teachers were observed at different times in the afternoon according to what was feasible, given the class timetable for the week. A smart phone application called *Gymboss* was used to provide beeps via headphones to the researcher, to signal the end of an interval and the start of the next interval.

3.5.8.2.4. Analysis

From the observation data, an *off-task behaviour* percentage for the whole class was calculated by dividing the number of off-task and disruptive intervals from the total number of intervals observed, multiplied by 100. A *disruptive behaviour* only percentage was also calculated. Further details on how this data was analysed is provided in Section 4.2.3.

3.5.8.2.5. Inter-observer agreement checks during the study

IOA checks were conducted on 33.3% of observations, which surpassed the recommended 20% (Kratochwill et al., 2010; Yoder & Symons, 2010), and a Cohen's kappa (1988) was used to calculate agreement whilst controlling for agreement acquired by chance. IOA checks were conducted with four PAs; one per observation. Each assistant received training from the researcher and was afforded the opportunity to practise using the schedule in the classroom.

Eleven IOA checks were conducted for on- and off-task behaviour, which ranged from 87.5% to 100% agreement, with a mean of 93.1%. Cohen's kappa ranged from 0.65-1.0. These results suggest that the observation schedule and coding used was highly reliable. IOA was only collected eight times for disruptive behaviour because disruptive behaviour was tallied separately during Time 1 data collection. This made it impossible to later ascertain which intervals were disruptive, for comparison. Data were recorded differently at Time 2 and 3 so that it was possible to do IOA checks for disruptive behaviour. IOA was also not calculated where PAs had forgotten to mark disruptive intervals as such and felt that their data was therefore invalid. On these occasions, the PAs went on to note disruptive behaviours accurately on their sheet, suggesting that the opportunity for further practise was useful. As recommended, discrepancy discussions were conducted to identify any issues with the behaviour categories and reduce observer drift (Yoder & Symons, 2010).

3.5.9. **Issues of data quality**

3.5.9.1. Reliability

Robson (2011) outlined four threats to the reliability of data obtained in research. These are considered in Table 3.4 alongside what was done to minimise these risks.

Threats	Description	Comment
Participant error	The pupils' on-task, off-task and disruptive behaviour might fluctuate according to the day / lesson.	Observation data was collected on different days to minimise day factors. Classes were also observed in a range of afternoon lessons.
	Teachers may be biased in their questionnaire responses due to tiredness, wanting to please the researcher or other factors.	Teachers were given up to two weeks at each time point to complete the questionnaires and encouraged to do a few each day, to eliminate the risk of tiredness and hasty completion.
Participant bias	The pupils may have behaved differently due to the researcher's presence in the classroom.	The researcher conducted several initial visits, which may have helped the pupils to become accustomed to her presence in the classroom. The researcher made a point to scan the room frequently so as to not alert the pupils that they were being watched specifically. The researcher also stood at the back of the room to be less visible.
Observer error	The researcher may have missed particular behaviours, leading to an incorrect coding.	Observations were conducted for 30-minutes to reduce tiredness. IOA checks were also conducted to reduce observer error.

Threats	Description	Comment
Observer bias	The researcher may have coded behaviours in a biased way, in order to obtain a particular result.	IOA checks were conducted frequently, however it was not possible to do this on each occasion.

Table 3.4: A table to outline the threats to reliability in this study and steps taken to reduce them.

3.5.9.2. Internal validity

A study that has internal validity is one in which a cause and effect relationship can be established between the independent variable (classroom management approaches) and the dependent variable (engagement/behaviour), rather than any changes (or no change) in the dependent variable being attributable to extraneous variables (Robson, 2011). There exist a number of potential threats to internal validity in research. The ones which were pertinent to this research are outlined in Table 3.5, adapted from Robson (2011), Mertens (2005) and Shadish et al., (2002).

Threats to internal validity	Description	Comment
History	Changes in the environment during the study affect the results	Waitlist control groups were included to control for environmental changes. Also, the groups were all selected from the same school and year group, to increase the likelihood that environmental changes would be similar across the groups.
Maturation	Biological or psychological changes to the participants over time, which is unrelated to the study's approaches	Waitlist control groups were included to control for maturation effects. The groups were selected from the same year group to increase the likelihood that maturation changes would occur at the same time for all groups.
Selection-maturation interaction	The risk that the results are affected by the difference in maturation across the groups	
Instrumentation	Aspects of the way in which participants are measured is changed during the study	The SDQ remained the same throughout all time measures. On-task, off-task and disruptive coding was agreed through IOA and discussion prior to beginning data collection. Also, the same observer conducted all observations to increase consistency of measuring. IOA checks were also frequently conducted.
Statistical regression	If participants are chosen due to being atypical, later testing will indicate that they are less atypical due to regression to the mean	The SDQ data of the pupils was not differentially analysed in terms of high scorers or low scorers. The class was analysed as a whole.

Threats to internal validity	Description	Comment
Differential selection	Differences in results between control and experimental groups are due to characteristic differences in the groups, rather than the approaches	The pupils could not be randomly allocated, however the lessons that were targeted were in the afternoons, where the pupils were in mixed ability groups. This increased the likelihood of differences within and across groups being balanced out. Furthermore, all of the classes were chosen from the same year group and school, to control for differences in age, maturation, demographics and school practices. All classes were similar in characteristics.
Experimental mortality	Dropping out of participants during the study	As there were a high number of participants being observed and having questionnaires filled out about them (4 classes and over 100 pupils), the likelihood of mortality affecting the results was minimised.
Experimental treatment diffusion	When a group inadvertently receives and/or implements the treatment when they should not	Teachers received training for only their approaches. All teachers were asked to not share what they were implementing in their classes. This, they agreed to. Fidelity checks were also conducted to assess diffusion of treatment in all classes.
Compensatory equalization of treatments or compensatory rivalry by control group	The pressure for the control group to receive the treatment or create improvements in other ways	The waitlist control groups were informed that they would be offered the approaches at the end of the study. All teachers were asked whether they would be happy to be in any of the groups and they said they would be.
Resentful demoralization of control group	Control group may perform lower due to not being chosen for the treatments	

Table 3.5: A table to outline the threats to internal validity in this study and steps taken to reduce them.

It was not possible to eliminate all threats to internal validity, however the table above highlights the ways in which the threats were reduced where possible. Other threats to internal validity include the approaches not being implemented to fidelity or the approaches not being implemented for long enough, for improvements to be observed (Mertens, 2005). Fidelity checks were conducted and teachers were asked to keep logs in order to ensure that fidelity occurred.

3.5.9.3. Threats to external validity

External validity relates to how generalisable the findings from this study are, to other populations (Robson, 2011). These threats include the possibility that the findings may not be applicable to others outside the groups that were part of this research (selection), and the school in which the study was conducted (setting). Furthermore, there is the risk that historical experiences which may have occurred during this study impacted the findings of this study and that the constructs which were studied (categories of on- and off-task behaviour and disruption) may be specific only to this setting and group (Robson, 2011).

The risk of construct effects were reduced by taking categories of on- and off-task behaviour and disruption from a range of previous research and piloting these categories in a different school. Historical experiences may not be controlled for, however, reporting them as potential extraneous variables would help with interpreting the results with caution. Selection and setting threats are pertinent to this research, however the methods of this study have been described in sufficient detail to allow for replication or generalisation to similar settings and groups. Replication in different contexts would reduce the threat of external validity (Mertens, 2005; Robson, 2011).

3.5.10. Ethical considerations

Guidelines for conducting ethical research have been produced by the University of Nottingham (2013), the Health Professions Council (2008), and the British Psychological Society (2009, 2014). The researcher examined these guidelines and considered the ethical issues that were

pertinent to this research in the planning and implementation stages. Ethical approval from the School of Psychology Ethics Committee at the University of Nottingham was sought and granted on 20th April 2016 (see appendix 8.5).

Following initial observations, parent information sheets and consent forms were distributed (see appendix 8.9). Unfortunately, only a small numbers of consent forms were returned, despite considerable effort that included sending several copies of the sheets home, emailing out information about the research, arranging several meetings, encouraging the pupils to bring the forms back, meeting with parents in the playground on several days, placing bulletins in the school newsletter and informing parents that pupils would be entered into a prize draw if their consent forms were returned *regardless of whether consent was granted or not* (see appendix 8.9). Two weeks after the consent forms were first sent out, an appeal was made to the School of Psychology Ethics Committee to approve opt-out consent procedures as the small number of consent forms returned would have led to a highly compromised design, which would have affected the validity of the findings. This approval was granted on 10th October 2016 (see appendix 8.5).

The particular ethical issues for this research are discussed below.

3.5.10.1. Autonomy and dignity of persons

The head teacher and teachers were informed of their right to withdraw from the study at any point. Pupils were not able to withdraw, however, their parents reserved the right to have their child's individual data withdrawn at any point, up to six weeks after the completion of the study (see appendix 8.9).

Informed consent was sought from the Chair of Governors, the head teachers and the class teachers (appendices 8.7 and 8.8). Opt-out consent procedures from parents were followed. Information sheets were distributed and parents were invited to email the researcher, attend meetings (see appendix 8.9), or speak to the researcher in the playground after school, in order to ask any questions. Informed

consent was not sought from the pupils due to the risk that their knowledge of the research's true aims would change their behaviour in class. Instead, assent was sought from them, for the researcher to observe their learning in lessons.

School staff and parents were assured that the data collected would be stored securely, that it would not be available to anybody other than the researcher, examiner and university tutor (see appendix 8.8 and 8.9). Anonymity in reporting the findings was also guaranteed.

The pupils were unaware that questionnaires had been filled out on them. The data was, however, completely anonymised before being given to the researcher and parents were able to opt-out of their child's data being collected if they wished. Furthermore, the pupils, staff and parents were debriefed about the true aims of the study and what data was collected, once the study had ended.

3.5.10.2. Maximising benefit and minimising harm and risk

In order to maximise benefit, the waitlist control groups were offered the approaches at the end of the study, if either of the approaches were found to be of benefit. It was recognised that filling out up to 30 questionnaires at three different time points could be aversive to teachers and so the teachers were informed in advance of when they would be given questionnaires and were afforded up to two weeks to complete them. Prior to beginning the research, the researcher asked the teachers whether any of the pupils were considered to be in a vulnerable state or whether they themselves were. It was reported that nobody was in a vulnerable state. Investigating the engagement of pupils in lessons was considered to be potentially a sensitive subject for teachers and so to reduce harm, they were reminded that it is typical for all classes to show some level of disengagement. They were also reminded that the approaches under investigation may be successful in increasing engagement.

The procedures were not considered to be aversive or stressful, being whole-class approaches to classroom management. However, it was

recognised that interdependent group contingency in which the behaviour of individuals could affect the whole group, may introduce conflict among the pupils. To reduce this risk, the pupils were explicitly instructed to be supportive and role-played ways to do this. *Being supportive* was a rule that was implemented as part of the approaches to further reduce the risk of conflict.

4.Results

4.1. Introduction

The previous chapter outlined the design and methodology of the present study, to answer the main research question:

How effective is a) a whole-class self-management and b) an interdependent group contingency approach, in terms of improving overall behaviour in junior school classes?

The focus of this research is on evaluating the efficacy of these approaches in improving behaviour and engagement in the target lessons as well as in improving behaviour outside of the target lessons. Four sub-questions were developed and a brief overview of the hypotheses for these is presented here (see Sections 5.2.2.1, 5.2.3.1, 5.2.4.1, and 5.2.7.1 for more detailed hypotheses):

1) ***What impact does whole-class self-management have on off-task and disruptive behaviours, as well as on the general behaviour of the whole class, as compared to control groups?***

Pupils in the self-management group will show a decrease in off-task and disruptive behaviour, as well as a decrease in general behavioural problems, and hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2, over and above those in the waitlist control conditions.

2) ***What impact does interdependent group contingency have on off-task and disruptive behaviours, as well as on the general behaviour of the whole class, as compared to control groups?***

Pupils in the interdependent group contingency group will show a decrease in off-task and disruptive behaviour, as well as a decrease in general behavioural problems and hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2, over and above those in the waitlist control conditions.

3) ***Which approach (self-management or interdependent group contingency) is most effective in reducing off-task and disruptive behaviours, and in improving general behaviour?***

Pupils in the self-management condition will show equivalent reductions in off-task and disruptive behaviour to pupils in the interdependent group contingency condition, between Time 1 and Time 2.

Pupils in the self-management condition will show a greater decrease in behavioural problems and hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2, compared to pupils in the interdependent group contingency condition.

4) Is there an added benefit to combining self-management with interdependent group contingency, with regard to off-task and disruptive behaviours, as well as general behaviour?

Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show a reduction in off-task and disruptive behaviour, as well as a reduction in general behavioural problems and hyperactivity and concentration problems between Time 2 and Time 3, which is greater than those observed in the control conditions.

The current chapter will begin by discussing the strengths and limitations of particular methods for data analysis, before providing a rationale for the analysis procedures chosen. The observation and questionnaire data for each research question will then be presented, analysed and summarised.

4.2. Approach to data analysis

4.2.1. Statistical tests

The purpose of psychological research is to test research predictions (Greene & D'Oliveira, 2005). Statistical tests provide a way for such predictions to be tested, in order to decide whether to accept or reject the null hypotheses; that is that there is no difference between scores (Field, Miles, & Field, 2012; Greene & D'Oliveira, 2005).

In quasi-experimental research data are often analysed using statistical tests, which may be parametric or non-parametric, depending on whether the data meet certain assumptions (Pallant, 2007; Robson, 2011). Parametric tests such as t-tests or ANOVAs are argued to be more robust and efficient than non-parametric tests, meaning that they are able to detect significant differences with a smaller sample, however, this has been contested where good non-parametric tests have been found to be as efficient as parametric tests (Robson, 2011). For parametric tests to be used, the data must be measured at the ratio or interval level, the sample must be randomly obtained from the population, the groups must show equal variance in their scores (homogeneity of variance), observations must be independent of each other, and the obtained data must be normally distributed (Dancey & Reidy, 2007; Pallant, 2007). The assumption of independence of observations must also be met for non-parametric tests to be used (W. E. Martin & Bridgmon, 2012). This assumption means that individual scores collected (e.g. each pupil's score), should not influence each other (Houser, 2015).

Most real world research is unable to meet these assumptions (Pallant, 2007). As such, one may choose to transform the data to resemble a normal distribution, however doing this may deceptively distort the data (Gomm, 2008). This is also argued to change the hypothesis that one is testing and choosing the wrong transformation method could have negative consequences on the analysis (Field, 2005). Instead, non-parametric tests which compare the median rather than the mean can be used (Pallant, 2007; Robson, 2011). Comparing the median rather

than the mean means the analysis is less affected by extreme scores (Dancey & Reidy, 2007). Non-parametric tests make no assumptions of the distribution of scores and as such, are recommended if the data collected is non-normal (Robson, 2011). However, as pre-determined groups tend to be allocated to conditions in a quasi-experiment, rather than individuals being randomly allocated to a condition, there is a risk that the assumption of independence of observation may be violated; for instance pupils within the same classroom may be more similar to each other than randomly chosen pupils (Shadish et al., 2002), as they share a common teacher or classroom environment (Sheng, 2008). As such, one pupil may influence another, affecting each other's scores on a given measure. A violation in the assumption of independence may invalidate inferences made by statistical tests, by affecting the accuracy of the significance (or *p)* level of the statistical tests (Sheng, 2008; Weiner & Craighead, 2010). As such, the risk of a Type 1 or Type 2 error may be increased (see Section 4.2.1.1 below).

4.2.1.1. Power

In research, it is possible to make Type 1 or Type 2 errors (Pallant, 2007). A Type 1 error refers to detecting an effect where there is none, and a Type 2 error refers to not detecting an effect that *is* present (Pallant, 2007; Robson, 2011). Statistical tests require sufficient power to be able to correctly detect significant differences between conditions in research, otherwise a Type 2 error is increased (Pallant, 2007). A power of 0.8 (80% chance of correctly detecting an effect) is desirable (J. Cohen, 1988), however, power is affected by sample size such that small sample sizes increase a Type 2 risk (Mertens, 2005; Pallant, 2007). It is argued that in comparison to parametric tests, non-parametric tests are less powerful, however, this is only true if the data meet the assumptions of parametric test (Field, 2005). Furthermore, whilst much has been produced to support calculating power for parametric tests (Mumby, 2002), it is not possible to calculate the actual power of a non-parametric test (Agarwal, 2003). Stevens (2012) argued that where groups contained only twenty participants, the power of parametric tests would be around .33 (low power), and for

groups as large as 50, the power would be .70 (medium power), if the level of significance remained at p=.05.

4.2.2. **Visual inspection**

An alternative to using statistical tests, particularly where these tests might lack statistical power due to a small sample size, is visual inspection of the data (Kadzin, 2003). Visual inspection is popular in single case experimental design (SCED) research (Nock, Michel, & Photos, 2008), and was employed in a number of the studies discussed earlier as part of the systematic review (Chafouleas et al., 2012; Denune et al., 2015; Glynn et al., 1973; Hoff & Ervin, 2013). Kratochwill et al., (2010) outlined four steps for conducting a visual inspection. These are:

1) Documenting a predictable baseline pattern of data.
2) Examining the data within each phase of the study to assess within-phase patterns.
3) Comparing the data from each phase to assess whether the implementation of the intervention was associated with an 'effect'.
4) Integrating all visual information to determine whether an effect has been demonstrated at three different time points.

Six features to examine patterns within and between phases were also described; these involved analysing the trend, variability, level and degree of overlap of the data points across phases, as well as examining the immediacy of effect and consistency of data points (Kadzin, 2003; Kratochwill et al., 2010; Lane & Gast, 2014):

Trend refers to the slope of the line of best fit for the data points within a phase.

Level refers to the mean score within a phase.

Variability refers to the degree to which the data points deviate from the line of best fit (Kratochwill et al., 2010).

Degree of overlap refers to the number of data points from one phase that overlaps with the data points in the previous phase.

Immediacy of effect refers to how quickly effects are observed in the dependent variable following the implementation of the intervention.

Consistency of data points refers to the degree to which data points in the same phases (i.e. all of the intervention phases or all of the baseline phases) are similar or consistent.

Not all six features would, however, be appropriate for analysing the observation data in a quasi-experimental design. In particular it would not be possible to measure immediacy of effect as this would require data to be collected throughout the implementation of the approach, or consistency of data points as this would require the intervention to be withdrawn and possibly reinstated following a return to baseline (Barlow, Nock, & Hersen, 2009).

Yoder and Symons (2010) described the need to conduct several observations to derive a reliable mean score of the observed phenomenon to detect behaviour change, however, rather than simply calculating means, visual inspection allows other characteristics of the data to also be analysed, to more reliably identify a causal link. Visual inspection can graphically illustrate clear intervention effects, which figures produced in statistical tests cannot show (Kadzin, 2003). It has, however, been criticised for being too subjective with little agreement between researchers as to what the data indicates (Kadzin, 2003). Furthermore, it has been argued that visual inspection may risk a Type 2 error, where small changes depicted graphically may lead to conclusions that there is no effect of the intervention (Kadzin, 2003), where in reality, the difference is significant for the individual or group of participants (Kadzin, 2003; Mertens, 2005).

4.2.3. Analyses chosen for this research

4.2.3.1. Visual inspection of observation data

In order to measure the impact of the classroom approaches on behaviour *in the target lessons,* partial interval recording data was collected. These data produced 'one off-task behaviour percentage' and a separate 'disruptive behaviour only percentage' for the entire class, per observation. As such, statistical tests could not be used to analyse this data. Instead, visual inspection was employed as it can graphically depict small changes in the behaviour of a limited sample and provides many ways of analysing the graphed data, beyond quoting changes in mean scores pre- and post.

Steps 1-3 of Kratochwill et al's (2010) four step procedure (see Section 4.2.2), were implemented. Step 4 could not be conducted as combining the data from several phases is specific to a design that returns to particular phases (i.e. returning to self-management only after implementing self-management with interdependent group contingency).

The level, trend, variability and overlapping data points of the observations were analysed, however, less emphasis was placed on trend in the analyses given that it was not possible to collect more than three observations at each time point. The trend data was therefore viewed with some caution.

Consistency of data points and immediacy of effect were not analysed as they are specific to phase designs employed within SCEDs.

4.2.3.1.1. Inter-rater reliability checks of graphed data

Due to concerns raised in the literature around the subjectivity of visual inspection (Kadzin, 2003), the researcher conducted inter-rater reliability checks on all of the graphed observation data presented in this section, with a Trainee EP who was familiar with visual inspection techniques. Both raters independently rated the change they believed

had occurred in relation to off-task behaviour and disruption (see appendix 8.22 for the sheet completed by each rater). A scale ranging from 1-6 was used to judge each graph, where 1=definite deterioration; 2=slight deterioration; 3=no change; 4=slight improvement; 5=definite improvement; and 6=unsure. The checks produced 93.3% agreement with a Cohen's kappa coefficient of 0.89 (almost perfect agreement) (Viera & Garrett, 2005).

4.2.3.2. Statistical analysis of Strengths and Difficulties Questionnaire data

The teacher version of the SDQ was used to measure general changes in behaviour following the implementation of the approaches. Only data from the conduct problems and hyperactivity and concentration subscales were analysed. Given the sample size for the questionnaire data (n=112 for Phase A and n=83 for Phase B), statistical analysis was chosen. Inspection of the descriptive statistics indicated that parametric tests would be unsuitable as the data had violated the assumptions of homogeneity of variance, normal distribution and random sampling (Dancey & Reidy, 2007). Therefore non-parametric tests were chosen (see section 4.2.3.2.1). The data were not transformed due to the risks associated with doing so (see Section 4.2.1).

4.2.3.2.1. Descriptive statistics

The descriptive statistics for the questionnaire data are presented in Table 4.1.

	Subscale	Class	Median	Mean	Variance	Standard Deviation	Skewness	Kurtosis
Time 1	Behavioural Problems	1 (SM)	1.00	0.83	0.97	0.99	**1.514**	**2.708**
		2 (IGC)	0.00	1.03	2.32	1.52	**1.696**	**2.824**
		3 (RR)	0.00	1.19	4.93	2.22	**2.415**	**5.447**
		4 (NC)	0.00	0.73	1.81	1.34	**1.608**	**1.166**
	Hyperactivity and concentration problems	1 (SM)	3.00	3.23	2.05	1.43	**1.149**	0.474
		2 (IGC)	3.00	3.14	5.62	2.37	0.601	-0.242
		3 (RR)	3.00	3.59	8.71	2.95	0.713	0.109
		4 (NC)	0.50	2.00	6.24	2.50	**1.017**	-0.042
Time 2	Behavioural Problems	1 (SM)	0.00	0.90	1.40	1.18	**1.139**	0.260
		2 (IGC)	2.00	2.52	7.69	2.77	0.898	-0.253
		3 (RR)	0.00	1.07	3.76	1.94	**2.482**	**6.191**
		4 (NC)	0.00	0.65	2.08	1.44	**2.231**	**3.861**
	Hyperactivity and concentration problems	1 (SM)	2.00	2.33	2.44	1.56	0.274	-0.416
		2 (IGC)	5.00	5.07	8.78	2.96	-0.178	**-1.059**
		3 (RR)	1.00	2.44	8.49	2.91	**1.044**	0.268
		4 (NC)	0.00	1.00	5.12	2.26	**2.581**	**6.329**

	Subscale	Class	Median	Mean	Variance	Standard Deviation	Skewness	Kurtosis
Time 3	Behavioural Problems	1 (SM + IGC)	0.50	0.87	1.15	1.07	0.997	-0.271
		2 (IGC + SM)	-	-	-	-	-	-
		3 (RR)	0.00	1.00	3.52	1.88	**2.24**	**4.70**
		4 (NC)	0.00	0.50	2.50	1.58	**3.49**	**12.40**
	Hyperactivity and concentration problems	1 (SM + IGC)	2.00	2.30	2.70	1.64	0.281	**-1.394**
		2 (IGC + SM)	-	-	-	-	-	-
		3 (RR)	1.50	2.46	9.06	3.01	**1.138**	0.390
		4 (NC)	0.00	0.77	5.23	2.29	**2.985**	**8.183**

Table 4.1: A table to present the descriptive statistics of the SDQ data.

Normality of the data is assessed by observing the skewness and kurtosis of the spread of data (Field, 2005; Pallant, 2007). If the skewness or kurtosis value falls outside of the range between 1 and -1, the distribution is considered non-normal (Dancey & Reidy, 2007). Table 4.1 highlights that all data at Time 1 was non-normal except for Classes 2 and 3 on the hyperactivity and concentration subscale. Normality of distribution can also be calculated using the Shapiro-Wilk statistic (Field, 2005). All Shapiro-Wilk statistics for the Time 1 data had a probability value of less than p=.05, indicating that all of the distributions violated the assumptions of normality (Field, 2005), required for parametric tests to be conducted.

Homogeneity of variance: Levene's Test of homogeneity of variance indicated that at Time 1, the variance of the classes did not differ significantly for behavioural problems (p=.075), however, did for hyperactivity and concentration problems (p=.012). This suggests that for hyperactivity and concentration problems at least, the data violated the assumptions of homogeneity of variance.

4.2.3.2.2. Non-parametric tests and significance levels

Given that the assumptions for parametric tests were violated, non-parametric tests were chosen to analyse the SDQ data.

A popular significance (alpha) level to use in social science research is p=.05 (Field, 2005), which states that there is only a 5% chance that a significant difference found within or between groups occurred due to chance. In order to answer all of the research questions, it was necessary to conduct twenty-three separate tests as there does not exist a non-parametric alternative to an ANOVA (Pallant, 2007). Multiple testing increases the probability of making a Type 1 error (Norman & Streiner, 2008), and so it is recommended that the alpha level is adjusted to reduce this risk (Field, 2005). The popular Bonferroni correction divides the probability level (p=.05) by the

number of tests (Norman & Streiner, 2008), however, it was not used in the present study as it is not recommended where more than 5 tests are conducted (Norman & Streiner, 2008). Instead, a more conservative p=.01 alpha level was adopted (Norman & Streiner, 2008), which was then checked against a less conservative correction; Holm's (1979) correction. The Holm procedure highlighted no difference to using a probability value of p=.01, in identifying statistical significance (see appendix 8.21 for an explanation of the Holm procedure and a table outlining the adjusted p level for each test using the Holm procedure).

Effect sizes as well as p levels were reported in the results as p levels alone are insufficient (Sullivan & Feinn, 2012). Effect sizes report the magnitude of the difference such that one could judge whether a 'non-significant' difference may have a large enough effect to potentially indicate clinical significance (Mertens, 2005). Effect sizes are reported as Pearson's correlation (r statistic) in non-parametric tests (Pallant, 2007). An r statistic of .1 indicates a small effect, whereas .3 indicates a medium effect and .5 indicates a large effect (J. Cohen, 1988).

The next section will outline and describe the findings for each research question in turn.

4.3. **Phase A results**

Research Question 1 was:

What impact does whole-class self-management have on off-task and disruptive behaviours, as well as on the general behaviour of the whole class, as compared to control groups?

This question explored the impact of self-management within lessons by collecting observation data at Time 1 and Time 2. The question also explored whether implementing self-management led to general changes in behaviour by analysing the *behavioural problems* and *hyperactivity and concentration problems* subscales of the SDQ.

Research Question 2 was:

What impact does interdependent group contingency have on off-task and disruptive behaviours, as well as on the general behaviour of the whole class, as compared to control groups?

This question explored the impact of interdependent group contingency within lessons by collecting observation data at Time 1 and Time 2. The question also explored whether implementing interdependent group contingency led to generalised changes behaviour by analysing the *behavioural problems* and *hyperactivity and concentration problems* subscales of the SDQ.

Research Question 3 was:

Which approach (self-management or interdependent group contingency) is most effective in reducing off-task and disruptive behaviours, and in improving general behaviour?

This question explored which approach (self-management or interdependent group contingency) led to greater improvements in off-task and disruptive behaviours within lessons by collecting observation data at Time 1 and Time 2. The question also explored which approach led to greater generalised changes in behaviour between Time 1 and Time 2, by analysing the *behavioural problems* and *hyperactivity and concentration problems* subscales of the SDQ.

The observation data for research questions 1, 2 and 3 (Phase A) is presented first, followed by the SDQ data. After this, the findings for research question 4 (Phase B) are presented.

4.3.1. Observation data

Each class' observation data for research questions 1, 2 and 3 (Phase A) is presented and interpreted individually before the data are combined onto one graph, to allow visual comparisons to be made with ease.

4.3.1.1. Class 1 (self-management)

4.3.1.1.1. Implementation of the approach

Unfortunately, logs of when the approaches were and were not implemented were not completed by the teacher. It was reported that self-management could only be conducted once or twice a week due to the nature of the activities or lessons on the other afternoons that rendered self-management unviable (e.g. planning, preparation and assessment time where the teacher was out of class or P.E.). Self-management was also not run during the four days that the teacher was absent. Dated rating sheets suggested that the approach was implemented in a total of eight lessons and of those, four were conducted during the two weeks of Time 2 data collection.

4.3.1.1.2. Off-task behaviour

Figure 4.1 shows the percentage of intervals where off-task behaviour was observed at Time 1, before the approach was implemented, and at Time 2, after four weeks. The vertical red line on the graph indicates the split between Time 1 and Time 2 data; this is where Phase A was implemented for four school weeks.

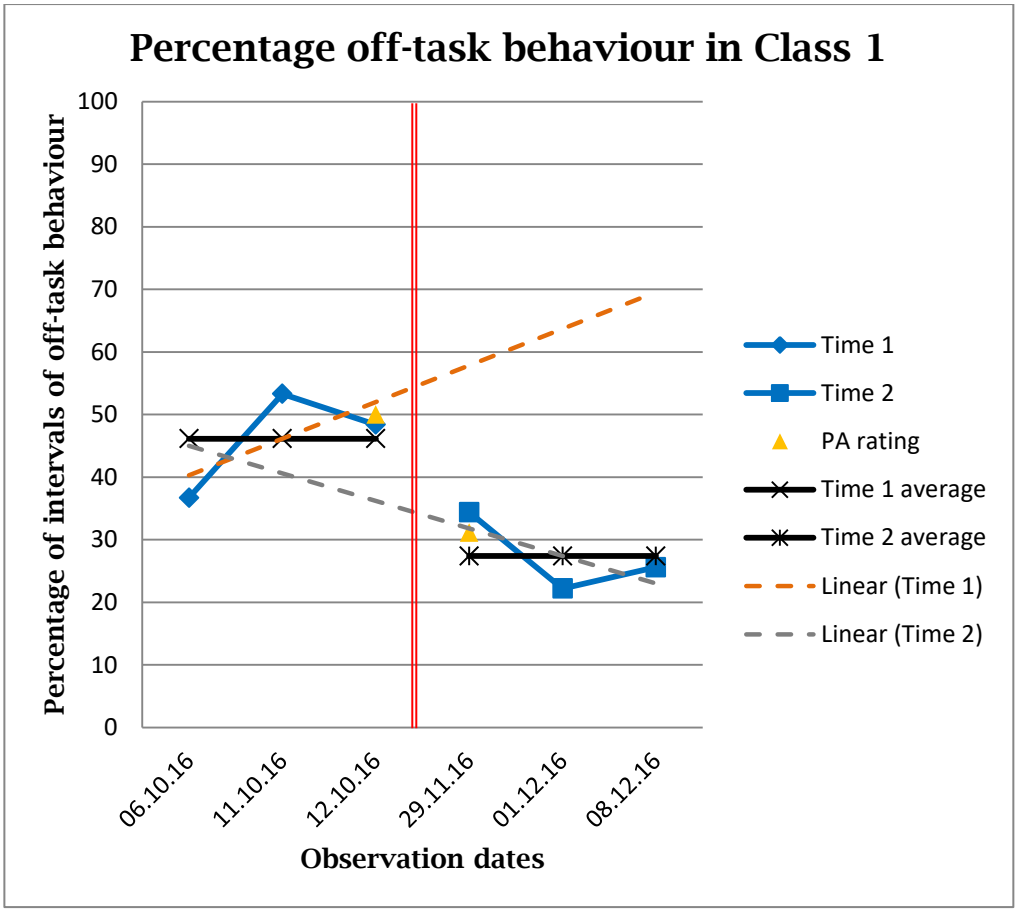


Figure 4.1: A graph to show the percentage of off-task intervals observed in Class 1, at Time 1 and Time 2.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained off-task behaviour at Time 1 was 46.13%. The mean percentage at Time 2 was reduced to 27.4%. Mean change = 18.73%.
Trend	The Time 1 data showed an upward trend with a slope of 5.85. The Time 2 data showed a downward trend with a slope of -4.4. Slope change = 10.25.
Variability	Scores at Time 1 were slightly more variable, with a range of 16.6% and standard deviation of 8.53%. At Time 2, there was less variability, with a range of 12.2% and standard deviation of 6.3%.
Overlapping data points	There were no overlapping data points between Time 1 and Time 2. Percentage of overlapping data points = 0%.

Table 4.2: A table to outline the findings from the visual inspection, on Class 1's graphed off-task behaviour data at Time 1 and Time 2.

The findings suggest off-task behaviour reduced after self-management was implemented. There was an almost 19% average decrease in the mean percentage of intervals where off-task behaviour was observed. Furthermore, the direction of the trend from Time 1 to Time 2 changed, such that an increasing trend for off-task behaviour at Time 1 became a decreasing trend at Time 2. There was some variability in the data, which may suggest that extraneous factors had some effect on the scores obtained, however this was not large. The absence of overlapping data points also suggests that there was a distinct change in off-task behaviour between Time 1 and Time 2 data. Taken together, the results of the visual inspection suggest a causal link between the implementation of self-management and a reduction in off-task behaviour.

4.3.1.1.3. Disruptive behaviour

Figure 4.2 shows the percentage of intervals where disruptive behaviours were observed at Time 1, before the approach was implemented, and at Time 2, after four weeks.

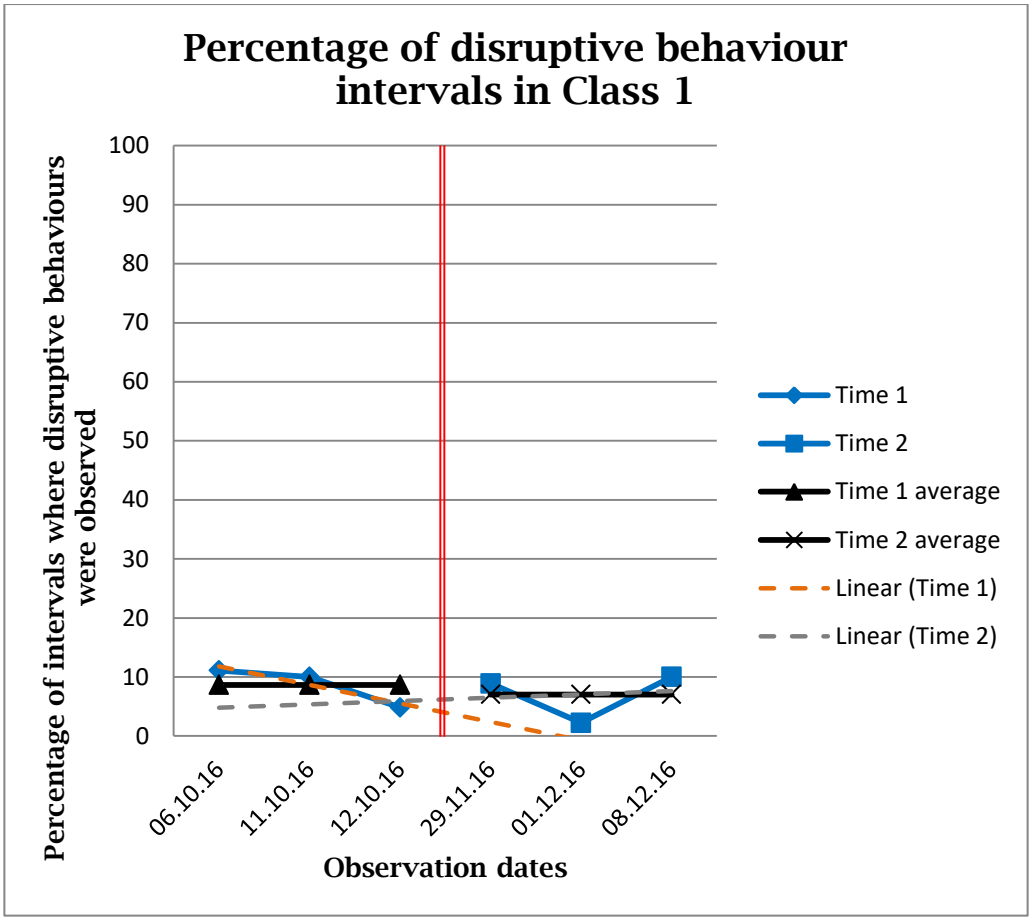


Figure 4.2: A graph to show the percentage of disruptive intervals observed in Class 1, at Time 1 and Time 2.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained disruptive behaviour at Time 1 was 8.65%. The mean percentage at Time 2 was slightly reduced to 7.04%. Mean change = 1.61%.
Trend	The Time 1 data showed a downward trend with a slope of -3.14. The Time 2 data showed a shallow upward trend, with a steeper slope of 0.56. Slope change = 3.7
Variability	Variability between Time 1 and Time 2 appeared unchanged. Scores at Time 1 had a range of 6.27% and standard deviation of 3.35%. At Time 2, the range was 7.78% with a standard deviation of 4.21%.
Overlapping data points	2 data point from Time 2 overlapped with Time 1 data points. Percentage of overlapping data points = 66.66%

Table 4.3: A table to outline the findings from the visual inspection, on Class 1's graphed disruptive behaviour data at Time 1 and Time 2.

The findings suggest self-management had no impact on disruptive behaviour. The difference in level between Time 1 and Time 2 for intervals of disruptive behaviour was less than 2%. Also, the majority of data points from Time 2 overlapped with Time 1 points, however initial rates of disruption were low. There was a negligible change in trend from a shallow decreasing trend to a shallow increasing trend. Little variability in the scores, suggested that the occurrence of disruptive behaviours was stable. Overall, the data suggested no real difference between Time 1 and Time 2 data points.

4.3.1.2. Class 2 (interdependent group contingency)

4.3.1.2.1. Implementation of the approach

The teacher in the interdependent group contingency class reported running the approach four days a week, however, she did not complete the teacher log. The goal sheet in the classroom showed that the approach had been implemented and the class were close to meeting their target. In total, the approach was implemented for only three weeks before the Christmas holidays, due to teacher absence through illness. The teacher then was absent for over three weeks in the run up to Christmas, meaning her Time 2 data could not be collected until after Christmas. Interdependent group contingency continued to be run one week after Christmas, and then Time 2 data was collected. This delay meant that Class 2 began Phase B three school weeks after the other classes entered Phase B.

4.3.1.2.2. Off-task behaviour

Figure 4.3 shows the percentage of intervals where off-task behaviour was observed at Time 1, before the approach was implemented, and at Time 2, after four weeks.

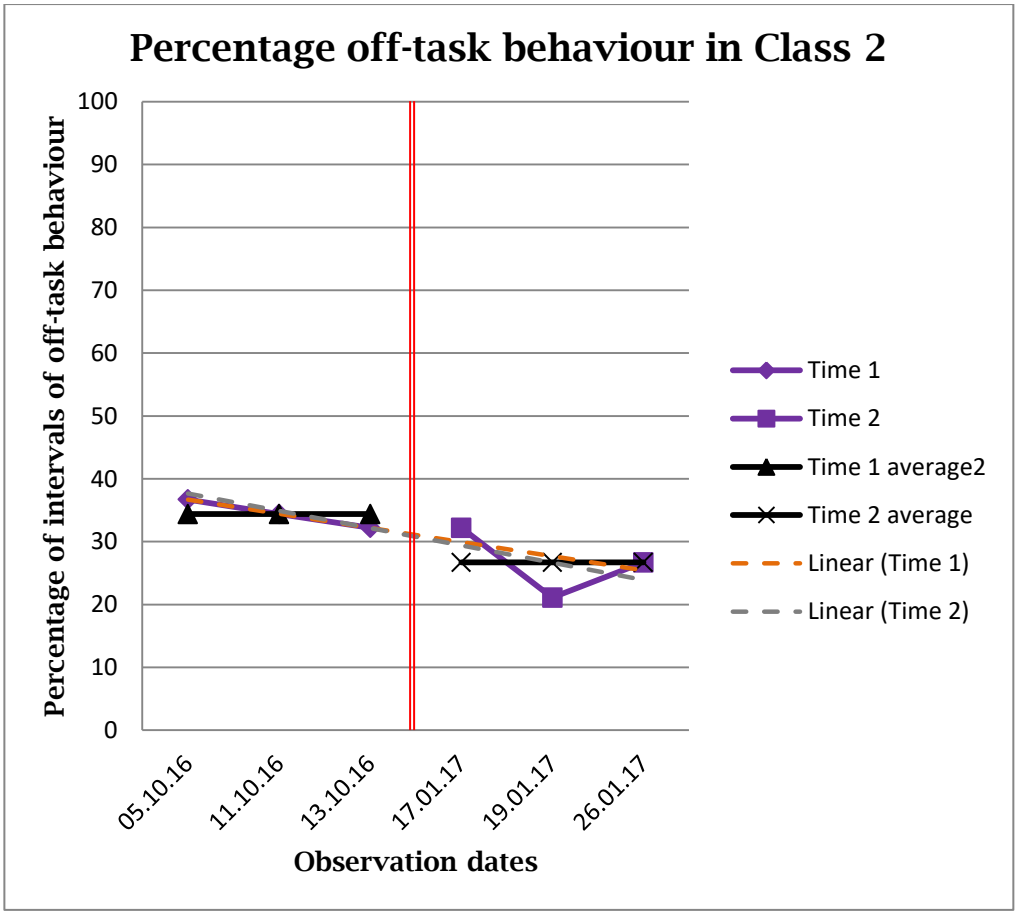


Figure 4.3: A graph to show the percentage of off-task intervals observed in Class 2, at Time 1 and Time 2.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained off-task behaviour at Time 1 was 34.4%. The mean percentage at Time 2 was reduced to 26.7%. Mean change = 7.7%.
Trend	The Time 1 data showed a downward trend with a slope of -2.25. The Time 2 data showed a similar downward trend with a slope of -2.75. Slope change = 0.5.
Variability	Scores at Time 1 were less variable, with a range of 4.5% and standard deviation of 2.25%. At Time 2, there was slightly more variability, with a range of 11.1% and standard deviation of 5.6%.
Overlapping data points	There was one overlapping data point between Time 1 and Time 2. Percentage of overlapping data points = 33.33%.

Table 4.4: A table to outline the findings from the visual inspection, on Class 2's graphed off-task behaviour data at Time 1 and Time 2.

4.3.1.2.3. Summary of findings: Off-task behaviour

The findings suggest off-task behaviour reduced slightly after interdependent group contingency was implemented. There was an almost 8% average decrease in the mean percentage of intervals where off-task behaviour was observed. The trend, however, did not change; at both time points, the trend suggested that off-task behaviour was decreasing before interdependent group contingency was implemented. There was little variability in the data, suggesting the scores were reliable. The small percentage of overlapping data points further suggests a change in off-task behaviour between Time 1 and Time 2 data. Overall, the results of the visual inspection suggest that there was a small reduction in off-task behaviour after interdependent group contingency was implemented.

4.3.1.2.4. Disruptive behaviour

Figure 4.4 shows the percentage of intervals where disruptive behaviours were observed at Time 1, before the approach was implemented, and at Time 2, after four weeks.

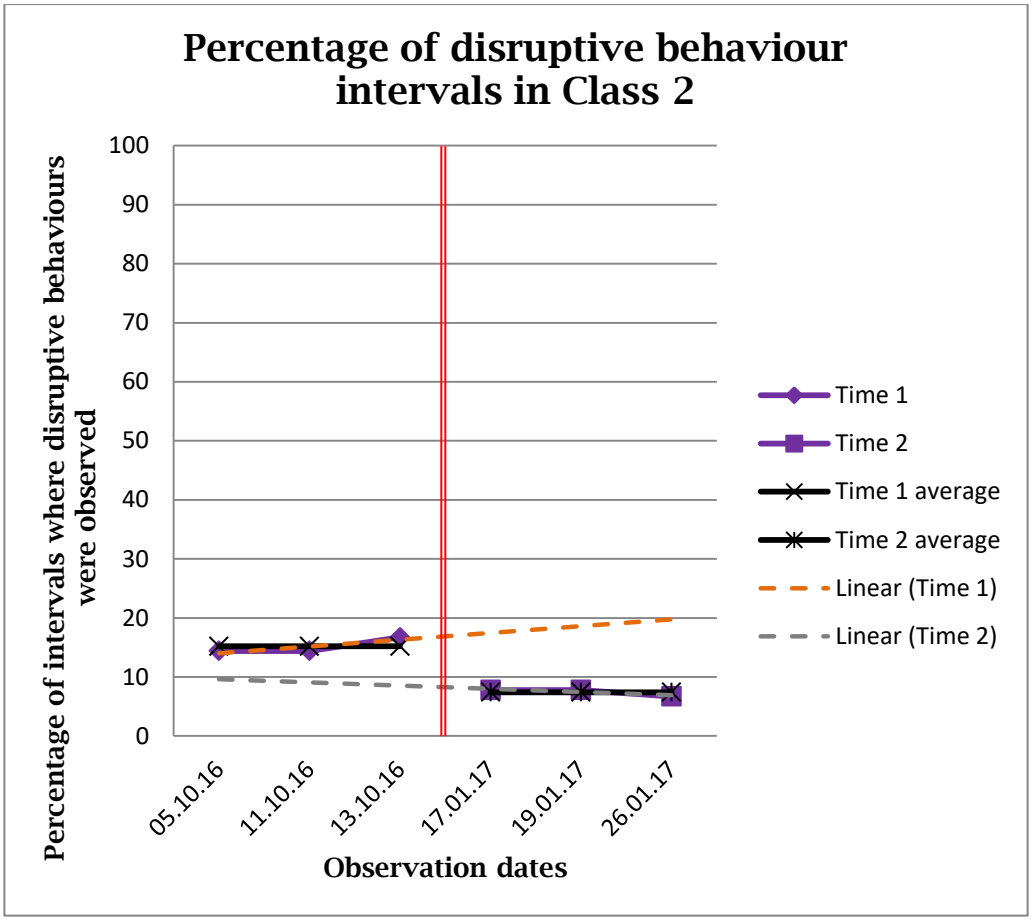


Figure 4.4: A graph to show the percentage of disruptive intervals observed in Class 2, at Time 1 and Time 2.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained disruptive behaviour at Time 1 was 15.2%. The mean percentage at Time 2 was slightly reduced to 7.4%. Mean change = 7.8%.
Trend	The Time 1 data showed a shallow upward trend with a slope of 1.15. The Time 2 data showed a shallow downward trend, with a slope of -0.55. Slope change = 1.7
Variability	Scores at both time points showed very little variability. Time 1 scores had a range of 2.3% and standard deviation of 1.33%. Time 2 scores had a range of 1.1% and standard deviation of 0.64%.
Overlapping data points	There were no overlapping data points between Time 1 and Time 2. Percentage of overlapping data points = 0%.

Table 4.5: A table to outline the findings from the visual inspection, on Class 2's graphed disruptive behaviour data at Time 1 and Time 2.

4.3.1.2.5. Summary of findings: Disruptive behaviour

The findings suggest that disruptive behaviour decreased following the implementation of interdependent group contingency. The data highlighted a mean change in intervals of disruptive behaviours of 7.8% between Time 1 and Time 2. There was little variability in the data, suggesting reliability of the measures. Furthermore, the trend between Time 1 and Time 2 changed from a shallow increasing trend to a shallow decreasing trend. There were also no overlapping data pints. Overall, the results suggest that interdependent group contingency produced a small but distinct decrease in disruptive behaviours.

4.3.1.3. Class 3 (rule reminder)

4.3.1.3.1. Implementation of the approach

The teacher reported that she had consistently started each afternoon lesson, four times a week, with a reminder of the rules. Nothing had been recorded in the teacher log book.

4.3.1.3.2. Off-task behaviour

Figure 4.5 shows the percentage of intervals where off-task behaviour was observed at Time 1, before the approach was implemented, and at Time 2, after four weeks.

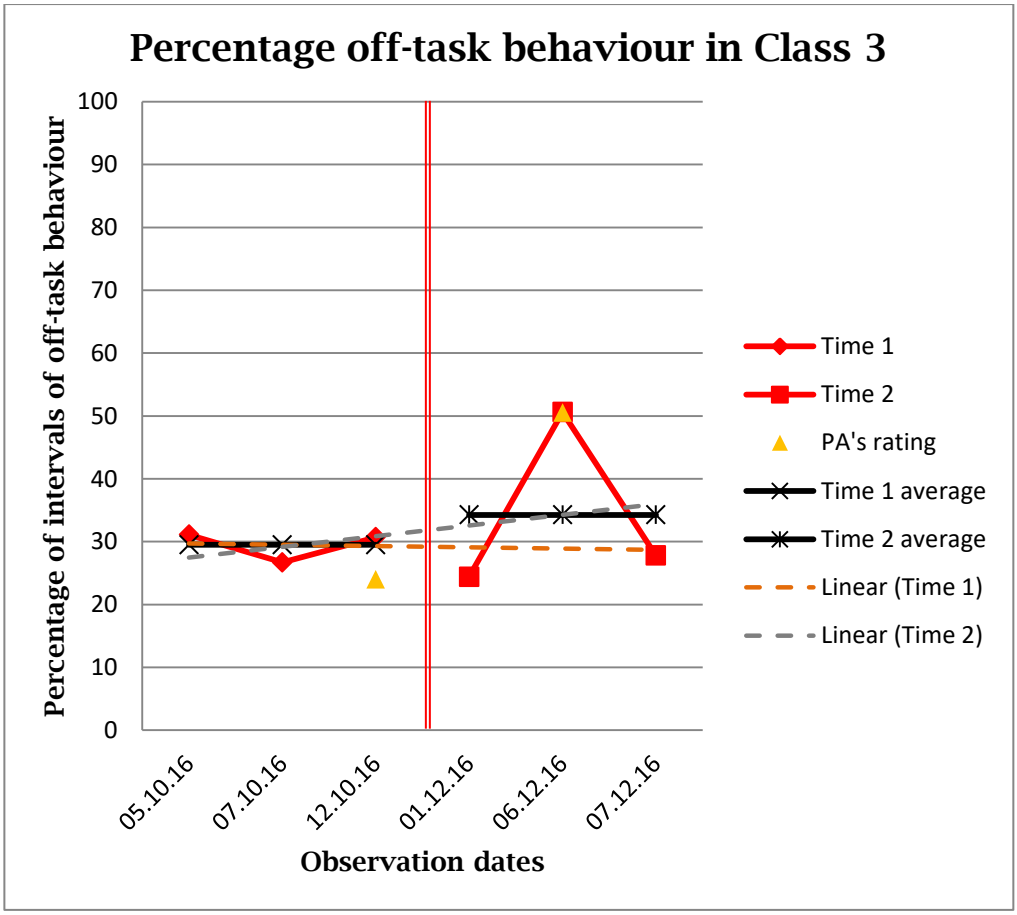


Figure 4.5: A graph to show the percentage of off-task intervals observed in Class 3, at Time 1 and Time 2.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained off-task behaviour at Time 1 was 29.5%. The mean percentage at Time 2 was slightly increased to 34.27%. Mean change = 4.77%.
Trend	The Time 1 data showed a slight downward trend with a slope of -0.2. The Time 2 data showed an upward trend, with a slope of 1.7. Slope change = 1.9.
Variability	Scores at Time 1 were less variable, with a range of 4.4% and standard deviation of 2.43%. At Time 2, there was much more variability, with a range of 26.2% and standard deviation of 14.25%.
Overlapping data points	1 data point from Time 2 overlapped with Time 1 data points. Percentage of overlapping data points = 33.33%

Table 4.6: A table to outline the findings from the visual inspection, on Class 3's graphed off-task behaviour data at Time 1 and Time 2.

Figure 4.5 highlights that the Time 1 (baseline) data points were stable and became highly variable at Time 2, however, the data point on 06.12.16 may have been an anomaly. This was the only observation that was conducted in the last half hour of the school day, which may have added an extraneous factor such as tiredness, which could have affected the score. Without this data point, the percentage of off-task behaviour appears more stable.

Figure 4.6 shows how the graph would look with the anomalous data point removed. This graph suggests less difference in level from Time 1 to Time 2. The data is less variable, and there is a greater proportion of overlapping data points.

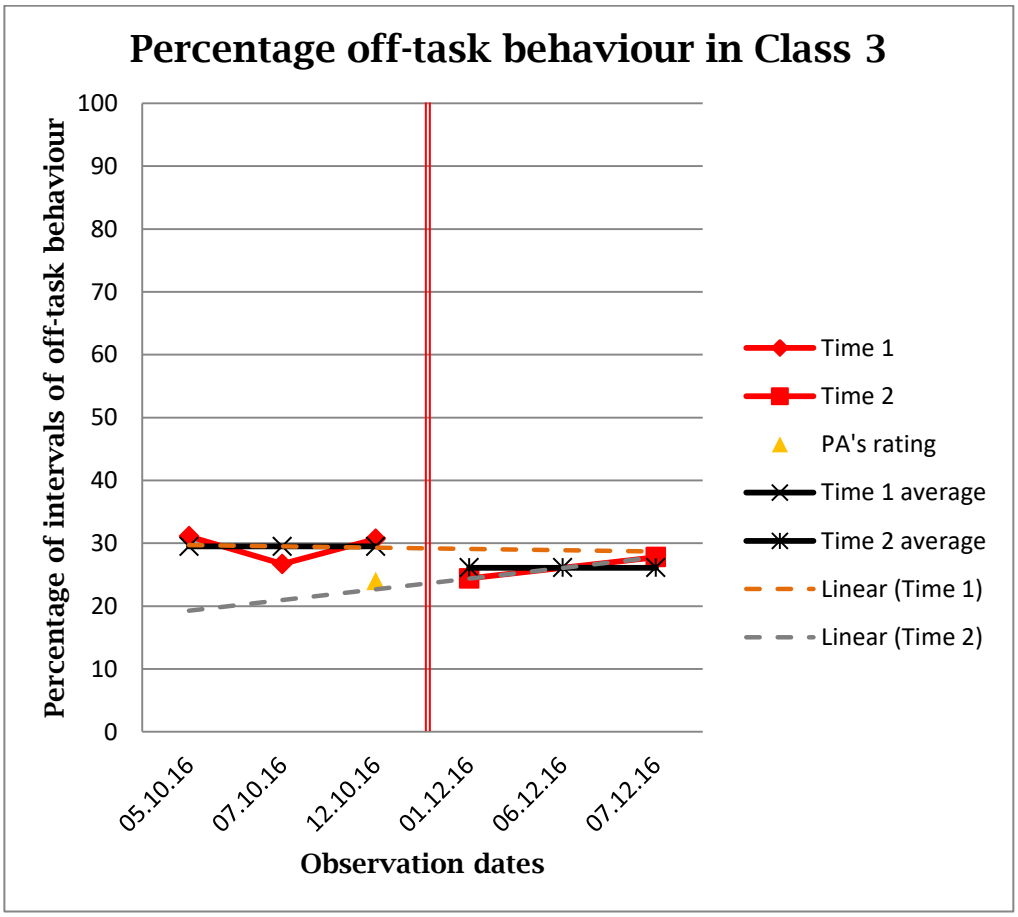


Figure 4.6: A graph to show the percentage of off-task intervals observed in Class 3, at Time 1 and Time 2, with the anomalous data point removed.

With or without the anomalous data point removed, the data from the rule reminder only class shows no real difference in off-task behaviour at Time 2, compared to Time 1. The change in level was negligible. There was a minimal change in trend which suggested slight increases

in off-task behaviour at Time 2, but not at Time 1. In both cases, there were overlapping data points.

4.3.1.3.3. Disruptive behaviour

Figure 4.7 shows the percentage of intervals where disruptive behaviour was observed at Time 1, before the approach was implemented, and at Time 2, after four weeks.

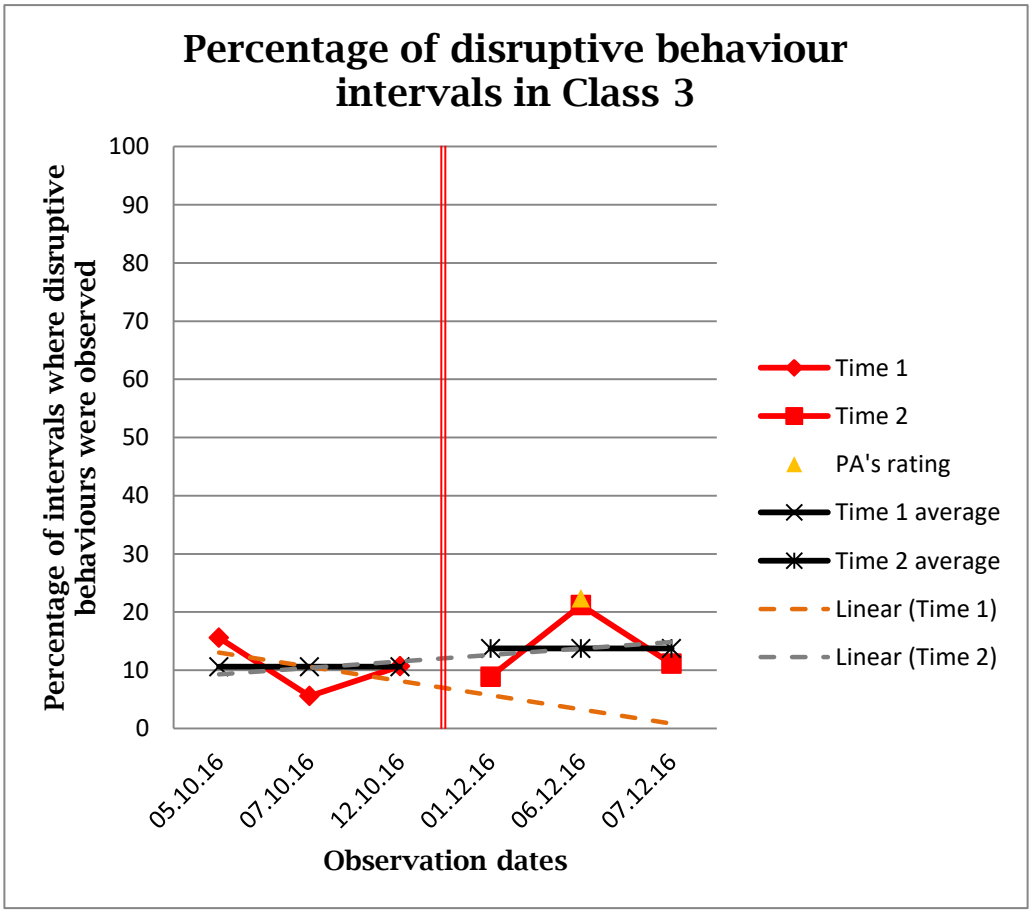


Figure 4.7: A graph to show the percentage of disruptive intervals observed in Class 3, at Time 1 and Time 2.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained disruptive behaviour at Time 1 was 10.6%. The mean percentage at Time 2 was slightly increased to 13.73%. Mean change = 3.13%.
Trend	The Time 1 data showed a downward trend with a slope of -2.45. The Time 2 data also showed an upward trend, with a shallow slope of 1.11. Slope change = 3.56
Variability	Scores at Time 1 were slightly less variable, with a range of 10% and standard deviation of 5%. At Time 2, there was a range of 12.29% and standard deviation of 6.55%.
Overlapping data points	2 data point from Time 2 overlapped with Time 1 data points. Percentage of overlapping data points = 66.66%

Table 4.7: A table to outline the findings from the visual inspection, on Class 3's graphed disruptive behaviour data at Time 1 and Time 2.

The results indicate that providing a rule reminder had no impact on disruptive behaviour. There were negligible changes to the average level of disruption and a large proportion of overlapping data points between Time 1 and Time 2. The change in trend indicated that there may actually be a slight increasing trend to disruption at Time 2, compared to the declining trend at Time 1.

4.3.1.4. Class 4 (no change)

4.3.1.4.1. Implementation of the approaches

From observing in this class and conducting a fidelity check, there was no evidence to suggest that there had been diffusion of treatment. The teacher reported being unaware of what the other classes were implementing and the other teachers reported that they had not shared with anybody else what they were doing in their classrooms for this research.

4.3.1.4.2. Off-task behaviour

Figure 4.8 shows the percentage of intervals where off-task behaviour was observed at Time 1, and at Time 2, after four weeks.

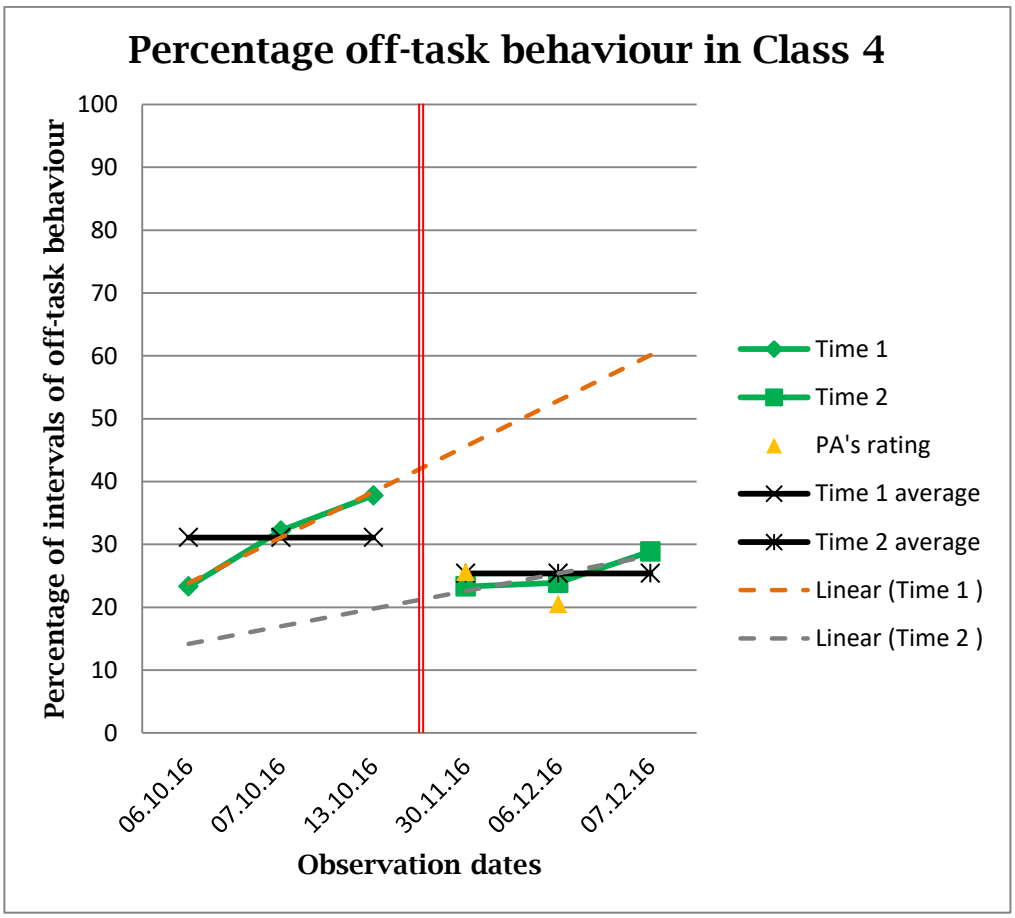


Figure 4.8: A graph to show the percentage of off-task intervals observed in Class 4, at Time 1 and Time 2.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained off-task behaviour at Time 1 was 31.1%. The mean percentage at Time 2 was slightly reduced to 25.4%. Mean change = 5.7%.
Trend	The Time 1 data showed an upward trend with a slope of 7.25. The Time 2 data also showed an upward trend, but with a less steep slope of 2.8. Slope change = 4.45.
Variability	Scores at Time 1 were more variable, with a range of 14.5% and standard deviation of 7.31%. At Time 2, there was less variability, with a range of 5.6% and standard deviation of 3.07%.
Overlapping data points	All three data points from Time 2 overlapped with Time 1 data points. Percentage of overlapping data points = 100%.

Table 4.8: A table to outline the findings from the visual inspection, on Class 4's graphed off-task behaviour data at Time 1 and Time 2.

The findings from the no change (waitlist control) class suggest that there was no difference in the class' off-task behaviour between Time 1 and Time 2. The visual inspection showed that there was a minimal decrease in level of off-task behaviour with all data points in Time 2 overlapping with Time 1 points. Furthermore, the data showed very little variability suggesting the scores obtained were stable and reliable. Additionally, there was no change in the direction of the trend, which at both time points was increasing.

4.3.1.4.3. Disruptive behaviour

Figure 4.9 shows the percentage of intervals where disruptive behaviour was observed at Time 1, and at Time 2, after four weeks.

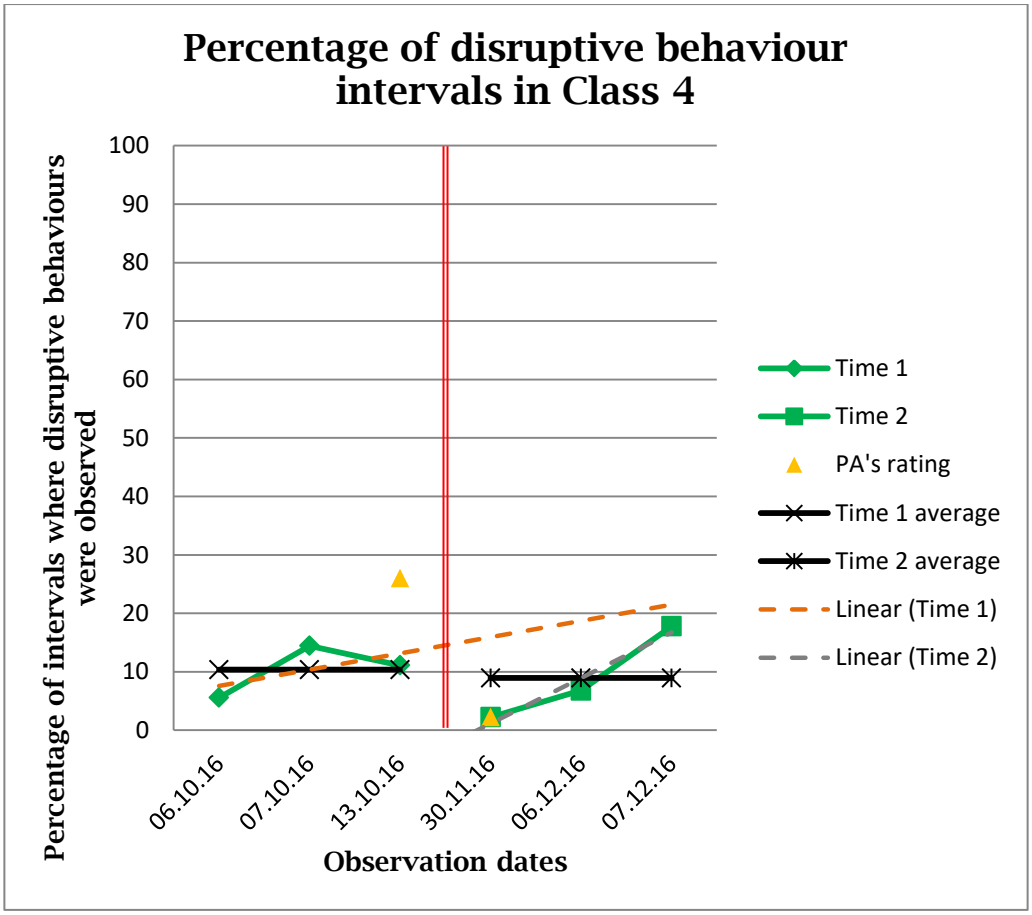


Figure 4.9: A graph to show the percentage of disruptive intervals observed in Class 4, at Time 1 and Time 2.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained disruptive behaviour at Time 1 was 10.37%. The mean percentage at Time 2 was slightly reduced to 8.93%. Mean change = 1.44%.
Trend	The Time 1 data showed an upward trend with a slope of 2.78. The Time 2 data also showed an upward trend, but with a steeper slope of 7.78. Slope change = 5
Variability	Scores at Time 1 were less variable, with a range of 8.89% and standard deviation of 4.49%. At Time 2, there was more variability, with a range of 15.56% and standard deviation of 8%.
Overlapping data points	1 data point from Time 2 overlapped with Time 1 data points. Percentage of overlapping data points = 33.33%

Table 4.9: A graph to show the percentage of disruptive intervals observed in Class 4, at Time 1 and Time 2.

The data on percentage of disruptive intervals suggested that there was no change in disruptive behaviour between Time 1 and Time 2. There was no change in level or trend direction, which for both time points suggested an increasing trend. Data points also overlapped. There was also little variability in the data points.

4.3.2. **Combined data**

Figure 4.10 and Figure 4.11 show the off-task and disruptive behaviours data from all classes on one graph, for ease of comparison.

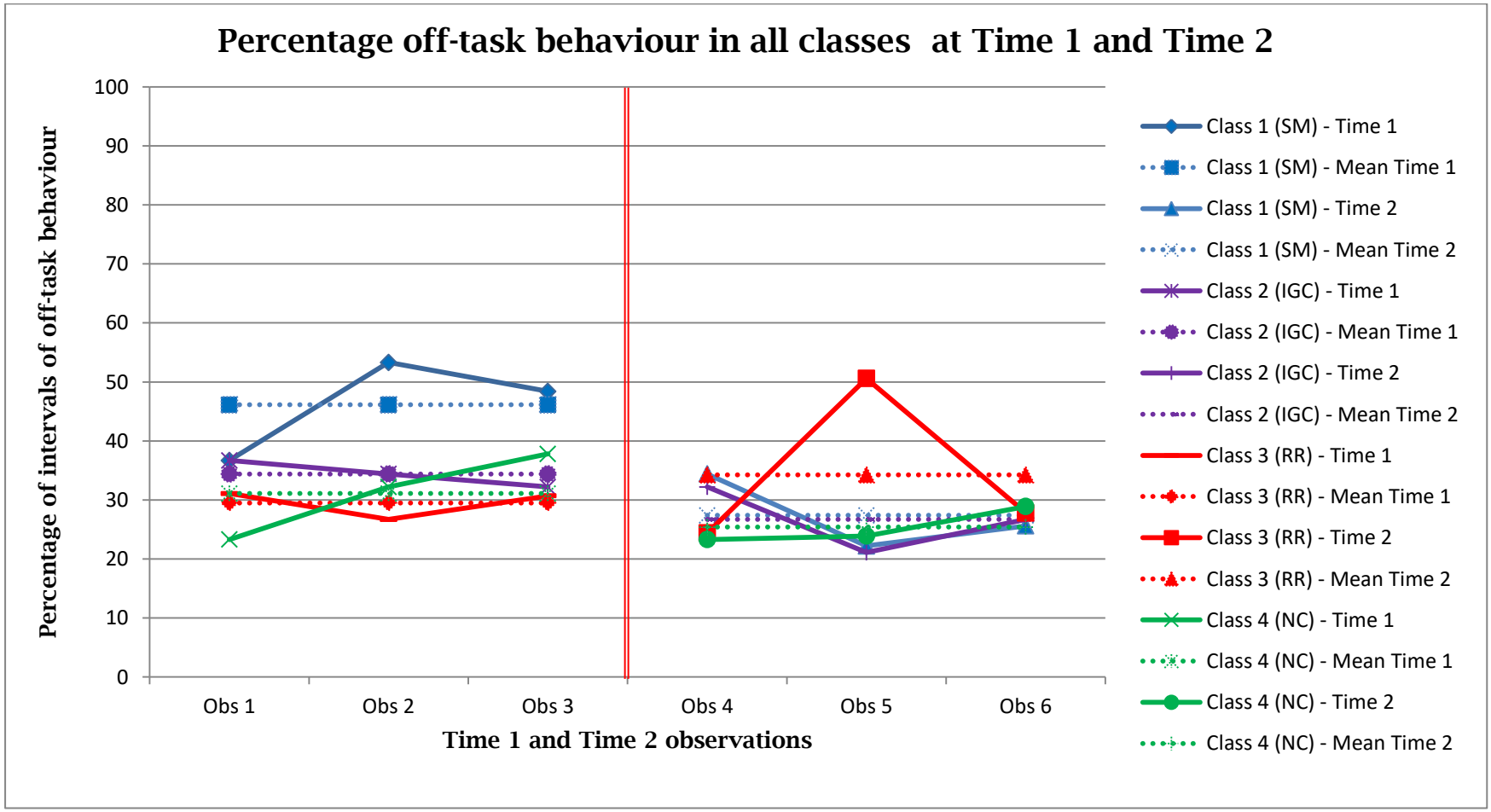


Figure 4.10: A graph to show the percentage of off-task intervals observed in all classes, at Time 1 and Time 2.

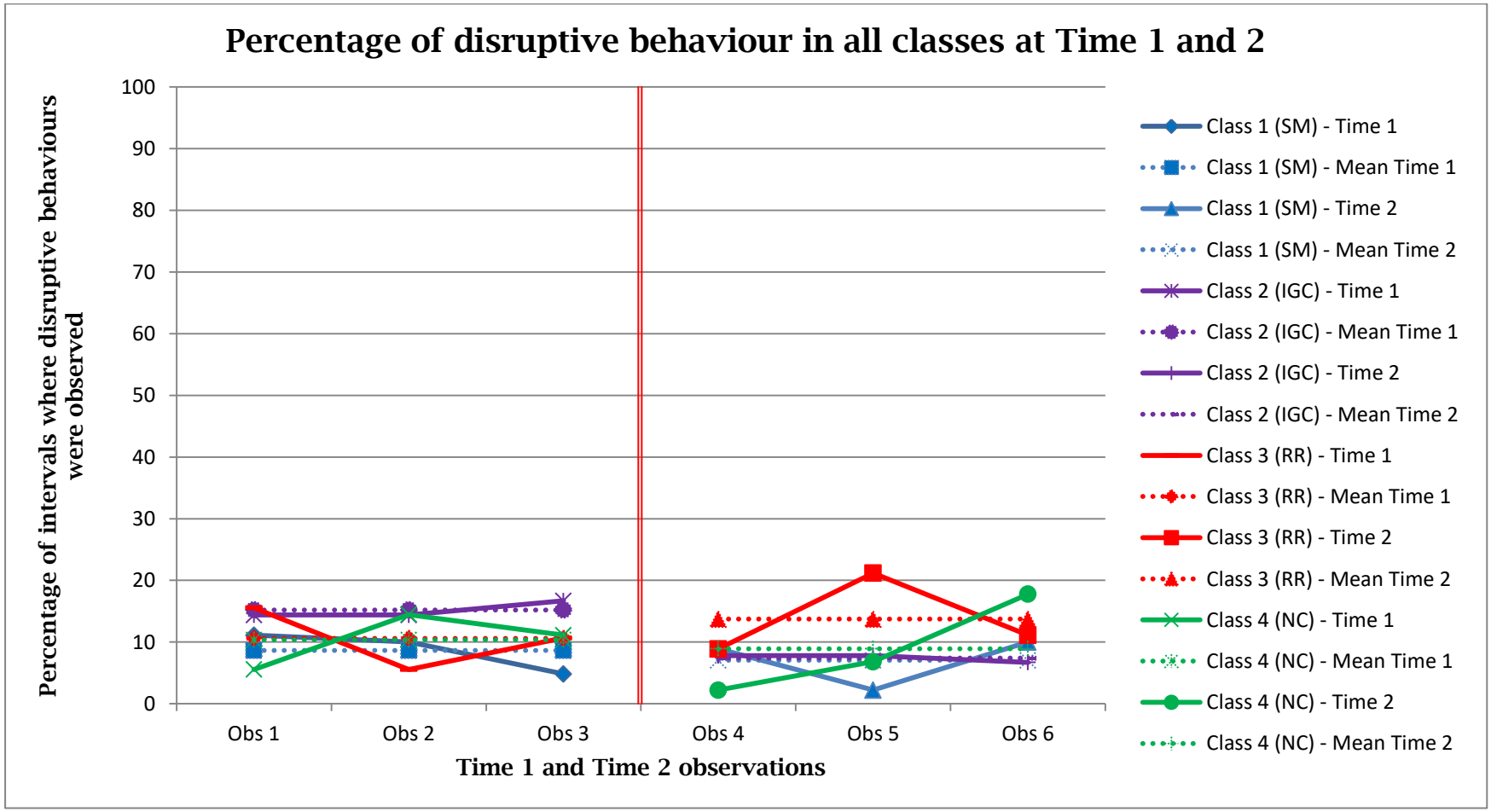


Figure 4.11: A graph to show the percentage of disruptive intervals observed in all classes, at Time 1 and Time 2.

4.3.3. Results summary of observation data for research question 1

Research Question 1: *What impact does whole-class self-management have on off-task and disruptive behaviours of the whole class compared to control groups?*

Overall, the results show that after self-management was implemented, there was a reduction in off-task behaviours, indicating that there may be a causal link and self-management may lead to less off-task behaviour in the target lessons. A change in off-task behaviour was not observed in the two control classes, suggesting that being reminded of the rules each day (rule reminder class) is insufficient on its own, to improve engagement. There was no difference in disruptive behaviour as a result of implementing self-management, in which Class 1 were equivalent to the other two classes at Time 2.

4.3.4. Results summary of observation data for research question 2

Research Question 2: *What impact does interdependent group contingency have on off-task and disruptive behaviours, as well as general behaviour of the whole class compared to control groups?*

Overall, the results show that after interdependent group contingency was implemented, there was a small reduction in off-task behaviours, indicating that there may be a causal link. This change was not observed in the control classes, suggesting that interdependent group contingency reduces off-task behaviour in the target lessons. Disruptive behaviours also reduced following interdependent group contingency, which was not the case for the control classes, suggesting this approach reduces disruption.

4.3.5. Results summary of observation data for research question 3

Research Question 3: *Which approach (self-management or interdependent group contingency) is most effective in reducing off-task and disruptive behaviours, as well as general behaviour issues?*

Overall, the results show suggest that self-management leads to greater reductions in off-task behaviour compared to interdependent group contingency, whereas it appears that only interdependent group contingency was effective in reducing disruptive behaviours.

The next section describes the findings from the SDQ for research questions 1, 2 and 3, which sought to measure generalised changes in behaviour.

4.3.6. Strengths and Difficulties Questionnaire data

The SDQ sought to measure whether there were generalised changes to the groups' behavioural problems and hyperactivity and concentration problems as a result of part-taking in the approaches. These results are presented in this section for research questions 1, 2 and 3 (Phase A).

Due to the use of multiple tests, a conservative alpha level of p=.01 was used, to reduce the risk of incorrectly accepting the null hypothesis (Norman & Streiner, 2008).

4.3.6.1. Behavioural problems

Figure 4.12 illustrates how all classes scored on behavioural problems at Time 1, before the approaches were implemented and at Time 2, after four school weeks.

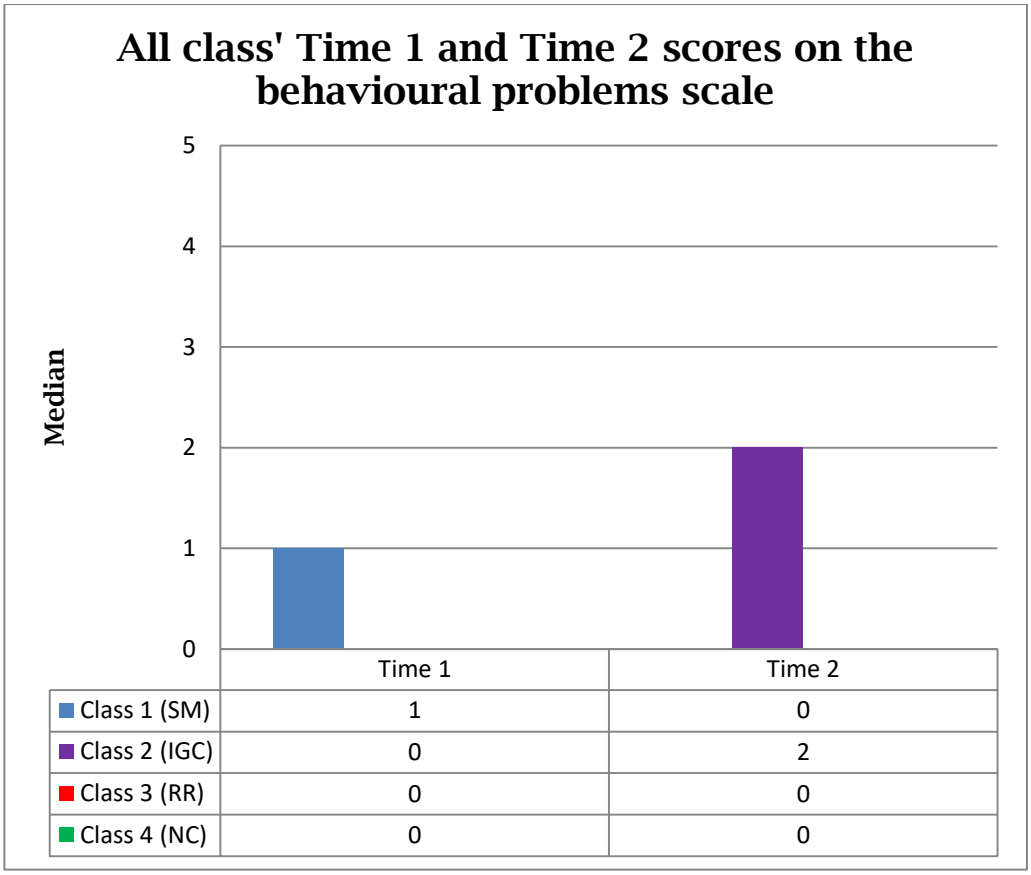


Figure 4.12: A graph to show all classes' median scores on the behavioural problems subscale of the SDQ, at Time 1 and Time 2.

A Kruskal-Wallis Test revealed that there was no statistically significant difference in behavioural problems across the four classes at Time 1 ($x^2(3)=2.142$, $p=.544$), or at Time 2 ($x^2(3)=11.21$, $p=.011$), however Figure 4.12 reveals that behavioural problems did increase between Time 1 and Time 2 in Class 2. The Wilcoxon Signed Rank Test indicated that there was no *statistically significant* difference in scores on behavioural problems between Time 1 and Time 2 in Class 1 ($z=-0.360$, $p=.719$, $r=.05$), suggesting that implementing self-management had no *significant* impact on behavioural problems. The difference between Time 1 and Time 2 in Class 2 (interdependent group contingency) *was,* found to be statistically significant however ($z=-3.455$, $p=.001$, $r=.45$). These results suggest that pupils in Class 2 scored significantly higher on behavioural problem measures after interdependent group contingency was implemented.

4.3.6.2. Hyperactivity and concentration problems

Figure 4.13 illustrates how all classes scored on hyperactivity and concentration problems at Time 1, before the approaches were implemented and at Time 2, after four school weeks.

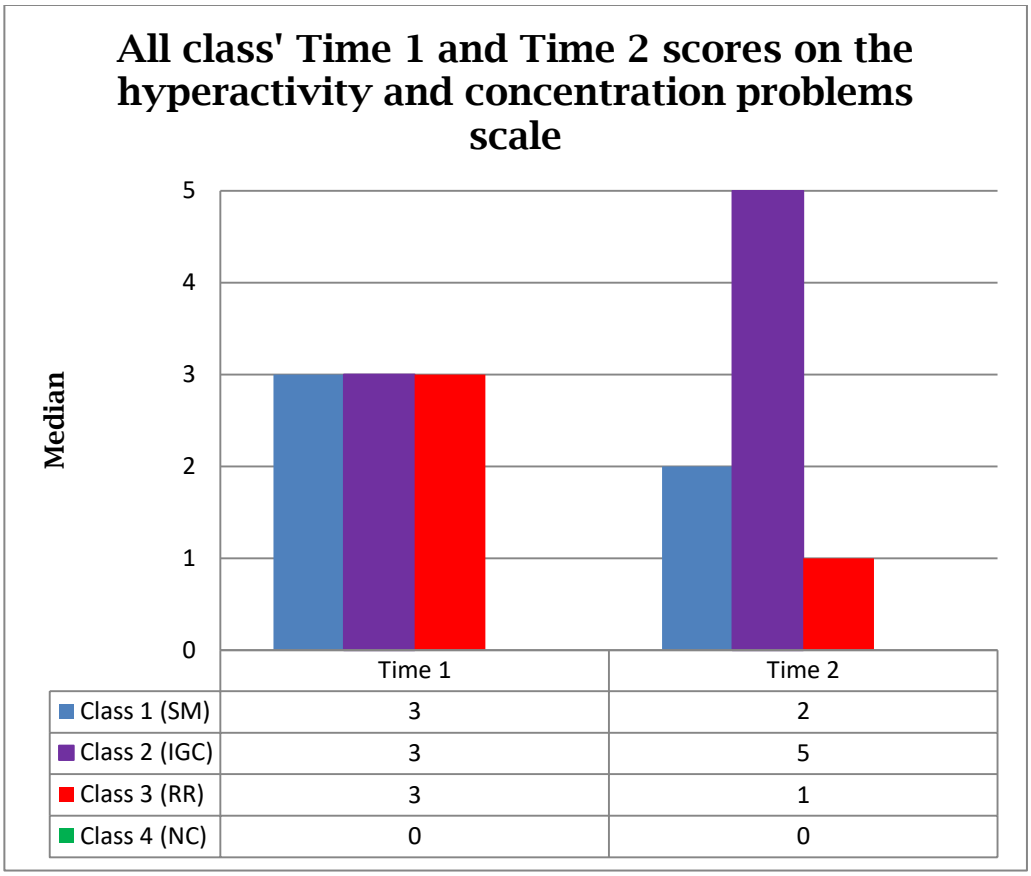


Figure 4.13: A graph to show all classes' median scores on the hyperactivity and concentration problems subscale of the SDQ, at Time 1 and Time 2.

A Kruskal-Wallis Test revealed that there was no statistically significant difference in hyperactivity and concentration problems across the four classes at Time 1 ($x^2(3)=7.507$, $p=.057$), however, the difference across the groups was significant at Time 2 ($x^2(3)=32.53$, $p=.00$). Mann-Whitney U Tests revealed that there was a significant difference between Class 1 (SM) and Class 2 (IGC), with Class 2 (IGC) showing significantly higher scores ($U=198.0$, $p=.00$, $r=.47$). There was also a

significant difference between Class 1 (SM) and Class 4 (NC), with Class 4 (NC) scoring significantly lower on hyperactivity and concentration problems (U=166.5, p=.00, r=.51). Class 2 (IGC) scored significantly higher compared to Class 3 (RR) (U=195.5, p=.001, r=.43), and Class 4 (U=95.0, p=.00, r=.66). There was no significant difference between Class 1 and Class 3 although a small effect size was detected (U=360.5, p=.469, r=.10). Figure 4.14 and Figure 4.15 illustrate the box plots for the data on hyperactivity and concentration problems at Time 1 and at Time 2.

The Wilcoxon Signed Rank Test indicated that Class 1 (SM) showed a significant decrease in hyperactivity and concentration scores between Time 1 and Time 2 (z=-3.287, p=.001, r=.42), although this was not significantly different to the control classes. Class 2 (interdependent group contingency) showed significantly higher scores in hyperactivity and concentration problems at Time 2 compared to Time 1 (z=-3.312, p=.001, r=.43).

Inspection of the box plots reveal that the self-management class and both control classes showed a reduction in their hyperactivity and concentration problems score at Time 2. This may indicate that the self-management class improved through maturation rather than as a result of the approach. The interdependent group contingency class was the only class to show higher scores at Time 2 for hyperactivity and concentration problems and a greater variance in scores than at Time 1.

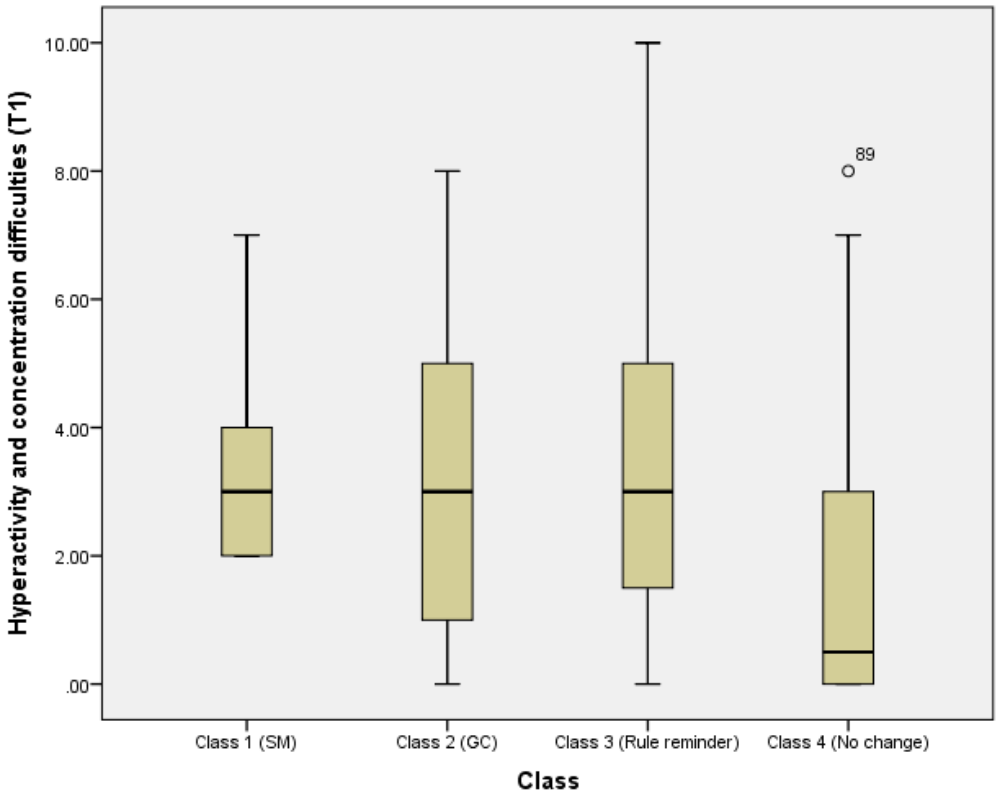


Figure 4.14: A boxplot to illustrate each class' spread of data on the hyperactivity and concentration problems subscale at Time 1.

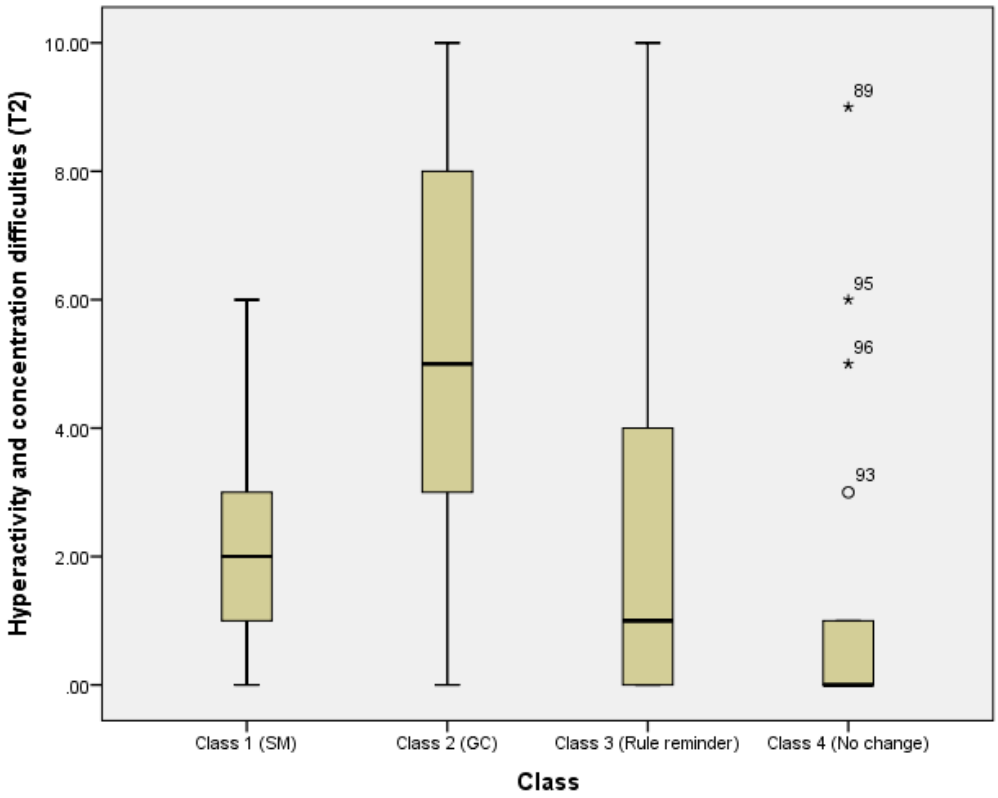


Figure 4.15: A boxplot to illustrate each class' spread of data on the hyperactivity and concentration problems subscale at Time 2.

4.3.6.3. Results summary of Strengths and Difficulties Questionnaire data for research question 1

Research question 1 aimed to ascertain whether implementing self-management would lead to generalised improvements in the behaviour of a whole class, through analysing the *behavioural problems* and *hyperactivity and concentration problems* subscales of the SDQ, as completed by class teachers.

The findings indicate that implementing self-management led to no significant generalised improvement in behavioural problems or hyperactivity and concentration problems, in comparison to the control groups. The results for behavioural problems however, should be interpreted with caution as initial scores were already at floor levels. This means that there was limited capacity for measuring improvement on this subscale.

4.3.6.4. Results summary of Strengths and Difficulties Questionnaire data for research question 2

Research question 2 aimed to ascertain whether implementing interdependent group contingency would lead to generalised improvements in the behaviour of a whole class, through analysing the *behavioural problems* and *hyperactivity and concentration problems* subscales of the SDQ, as completed by class teachers.

The findings indicate that the interdependent group contingency class scored significantly higher on hyperactivity and concentration problems than the control classes at Time 2, however, there was no significant difference between Class 2 and the control classes on behavioural problems at Time 1 or at Time 2. The lack of statistical significance between the groups on behavioural problems at Time 2 may be due to the more conservative p=.01 being adopted as the level of significance. Inspection of Figure 4.12 suggests there was a

noticeable difference, with Class 2 showing greater behavioural problems in comparison to the control groups.

4.3.6.5. Results summary of Strengths and Difficulties Questionnaire data for research question 3

Research question 3 sought to ascertain whether self-management or interdependent group contingency led to greater generalised improvements in behaviour of a whole class, through analysing the *behavioural problems* and *hyperactivity and concentration problems* subscales of the SDQ, as completed by class teachers.

The findings suggest that implementing self-management had no significant impact on behavioural problems, however, implementing interdependent group contingency led to significantly increased behavioural problems. Despite this, statistical analyses suggest that both groups did not differ significantly in their Time 1 or Time 2 scores, despite noticeable differences when comparing the medians of these groups (see Figure 4.12).

On the hyperactivity and concentration problems subscale, self-management appeared to lead to significantly lower scores, although this was no different to the control classes, whereas interdependent group contingency appeared to result in significantly higher scores. Although the groups were equivalent at Time 1, at Time 2, they appeared to be significantly different.

4.3.7. Overall summary of results for Phase A

Overall the research suggests that self-management reduces off-task behaviour more so than interdependent group contingency in the target lesson, but that only interdependent group contingency reduces disruption. These effects were not observed in the waitlist control conditions. The results also suggested that implementing self-management led to no significant change in general behavioural problems or hyperactivity and concentration problems compared to controls, whereas interdependent group contingency led to

significantly higher scores on hyperactivity and concentration problems, and noticeable but non-significantly higher scores on behavioural problems, in comparison to control groups.

4.4. Phase B results

Research Question 4 was:

Is there an added benefit to combining self-management with interdependent group contingency, with regard to off-task and disruptive behaviours, as well as general behaviour?

This question explored the impact of combining self-management with interdependent group contingency on off-task and disruptive behaviours within the target lessons by collecting observation data at Time 2 and Time 3. The question also explored whether combining these approaches led to generalised changes in behaviour by analysing the *behavioural problems* subscale and the *hyperactivity and concentration problems* subscale of the SDQ.

Unfortunately, due to long-term teacher absence and subsequent resignation, it was not possible to collect Time 3 observation or questionnaire data from Class 2. As such, this research question is answered using the data from Classes 1, 3 and 4 only.

4.4.1. Observation data

Each class' data is presented and interpreted individually before the data are combined onto one graph, to allow for ease of visual comparison.

4.4.1.1. Class 1 (interdependent group contingency added to self-management)

4.4.1.1.1. Implementation of the approach

The self-management rating sheets returned to the researcher suggest that the combined approach (self-management with added interdependent group contingency) appears only to have been

implemented six times. Of those six occasions, only three were implemented before Time 3 data collection commenced.

4.4.1.1.2. Off-task behaviour

Figure 4.16 shows the percentage of intervals where off-task behaviour was observed at Time 2, before interdependent group contingency was added, and at Time 3, after the combined approach had been implemented for four weeks. The vertical blue line on the graph indicates the split between Time 2 and Time 3 data; this is where Phase B was implemented for four school weeks.

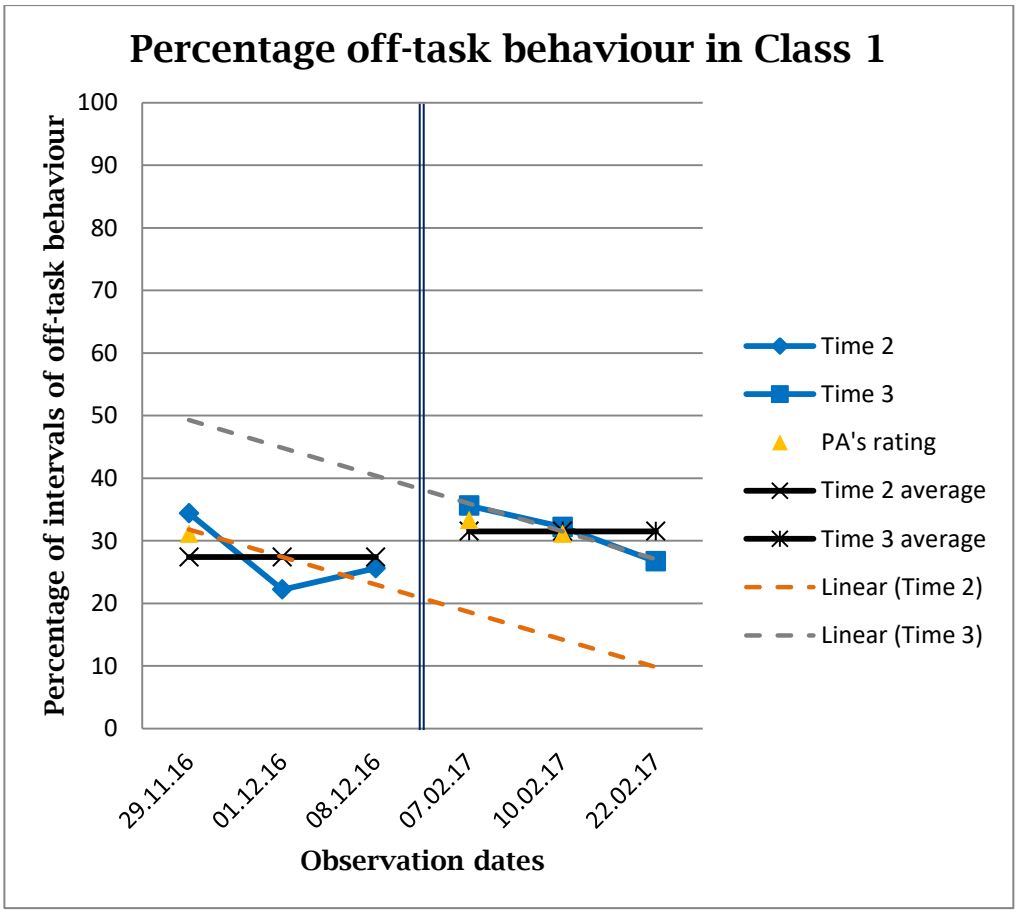


Figure 4.16: A graph to show the percentage of off-task intervals observed in Class 1, at Time 2 and Time 3.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained off-task behaviour at Time 2 was 27.4%. The mean percentage at Time 3 increased slightly to 31.5%. Mean change = 4.1%.
Trend	The Time 2 data showed downward trend with a slope of -4.4. The Time 3 data showed a similar trend with a slope of -4.45. Slope change = 0.05.
Variability	Scores at Time 2 were slightly more variable than at Time 3, with a range of 12.2% and standard deviation of 6.3%. At Time 3, there was slightly less variability, with a range of 8.9% and standard deviation of 4.49%.
Overlapping data points	2 data point from Time 3 overlapped with Time 2 data points. Percentage of overlapping data points = 66.66%

Table 4.10: A table to outline the findings from the visual inspection, on Class 1's graphed off-task behaviour data at Time 2 and Time 3.

These findings suggest that there was no change in off-task behaviour between Time 2 and Time 3, in Class 1. The data showed a negligible change in level, the same trend at both time points and a large amount of overlapping data points. Furthermore, there was little variability in the data, which suggests it is reliable. Overall, these findings suggest that adding group contingency to an already operating self-management approach does not result in further reductions in off-task behaviour.

4.4.1.1.3. Disruptive behaviour

Figure 4.17 shows the percentage of intervals where disruptive behaviour was observed at Time 2, before interdependent group contingency was added, and at Time 3, after the combined approach had been implemented for four weeks.

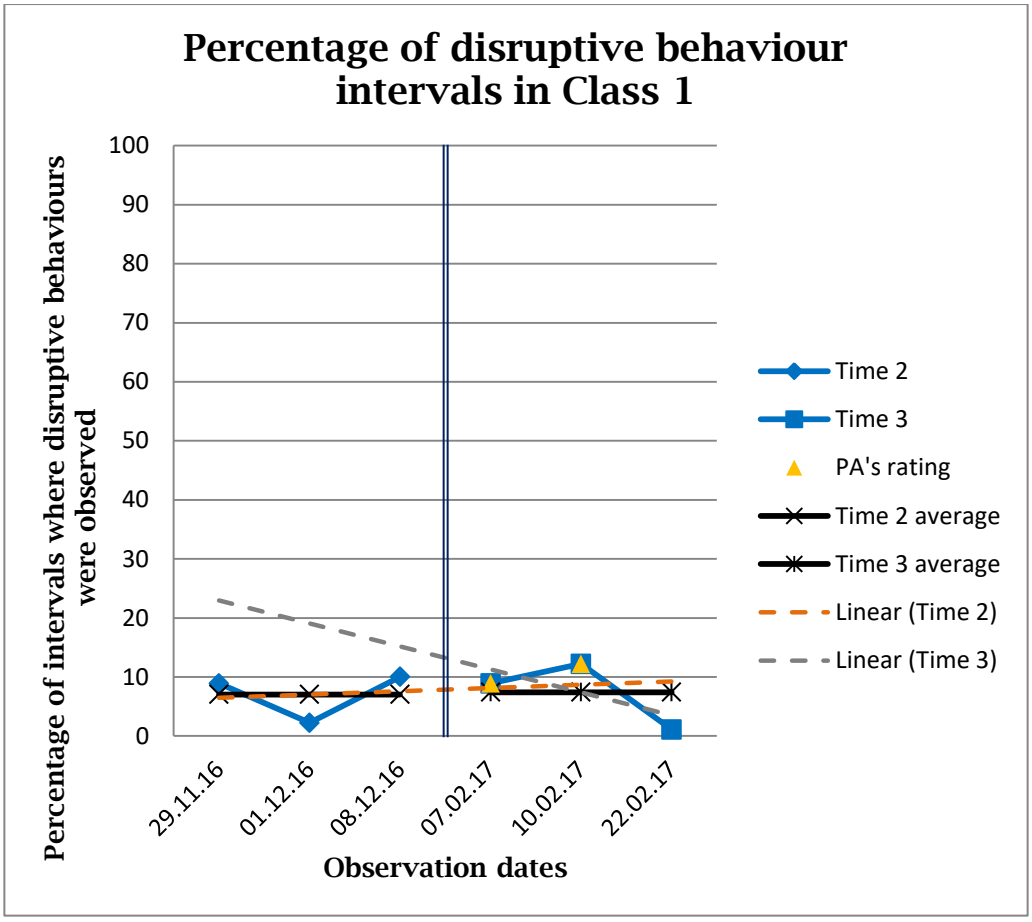


Figure 4.17: A graph to show the percentage of disruptive intervals observed in Class 1, at Time 2 and Time 3.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained disruptive behaviour at Time 2 was 7.04%. The mean percentage at Time 3 was 7.4%. Mean change = 0.36%.
Trend	The Time 2 data showed a slight upward trend with a slope of 0.56. The Time 3 data showed a steeper downward trend, with a slope of -3.9. Slope change = 4.46.
Variability	Scores at Time 2 (range = 7.78% and standard deviation = 4.21%) were slightly less variable than scores at Time 3 (range = 11.12% and standard deviation = 5.71%).
Overlapping data points	1 data point from Time 3 overlapped with Time 2 data points, however all Time 2 data points overlapped with Time 3 data points. Percentage of overlapping data points = 33.33%

Table 4.11: A graph to show the percentage of disruptive intervals observed in Class 1, at Time 2 and Time 3.

The results suggest that there was no change in disruption between Time 2 and Time 3 in Class 3. There was no change in the level and all of the Time 2 data points overlapped with the Time 3 data points. The trend did change substantially to a steep decreasing trend at Time 3, however the limited data points makes the trend a less reliable measure. Overall the findings suggest that adding group contingency to an already operating self-management approach leads to no change in the amount of disruptive behaviour.

4.4.1.2. Class 3 (rule reminder)

4.4.1.2.1. Implementation of the approach

The teacher reported that she had consistently started each afternoon lesson, four times a week, with a reminder of the rules. Nothing had been recorded in the teacher log book.

4.4.1.2.2. Off-task behaviour

Figure 4.18 shows the percentage of intervals where off-task behaviour was observed at Time 2 and after four weeks, at Time 3.

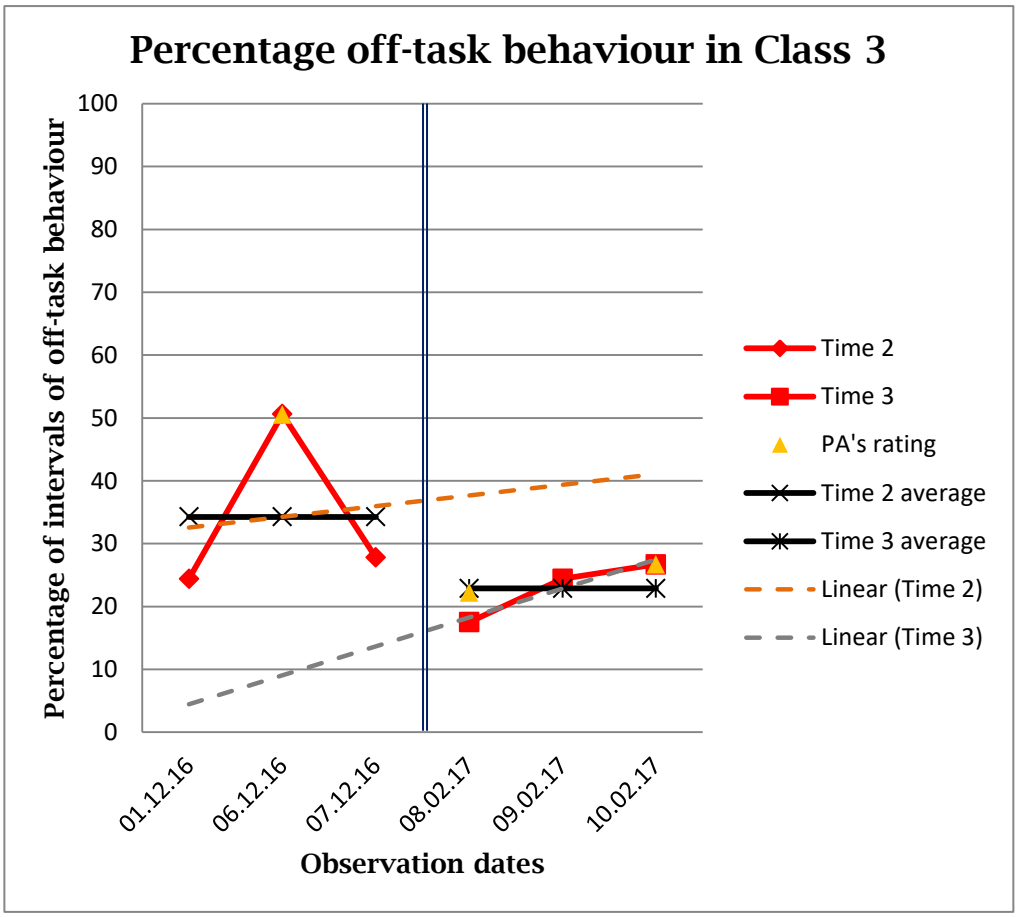


Figure 4.18: A graph to show the percentage of off-task intervals observed in Class 3, at Time 2 and Time 3.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained off-task behaviour at Time 2 was 34.27%. The mean percentage at Time 3 decreased to 22.87%. Mean change = 11.4%.
Trend	The Time 2 data showed a slight upward trend with a slope of 1.7. The Time 3 data also showed an upward trend, with a slope of 4.6. Slope change = 2.9.
Variability	Scores at Time 2 were more variable, with a range of 26.2% and standard deviation of 14.25%. At Time 3, there was much less variability, with a range of 9.2% and standard deviation of 4.79%.
Overlapping data points	2 data point from Time 3 overlapped with Time 2 data points. Percentage of overlapping data points = 66.66%

Table 4.12: A table to outline the findings from the visual inspection, on Class 3's graphed off-task behaviour data at Time 2 and Time 3.

The findings suggest that reminding the class of the rules daily led to a slight reduction in off-task behaviour between Time 2 and Time 3, in Class 3. There was a change in level of 11.4%, however, with the anomaly (06.12.16) removed, change in level was less (6.63%). Furthermore, the data at both time points suggested upward trends and the majority of data points at Time 3 overlapped with those at Time 2. Aside from the 06.12.16 data point, the data shows very little variability, suggesting that it is reliable. Overall, the results should very much be interpreted with caution.

4.4.1.2.3. Disruptive behaviour

Figure 4.19 shows the percentage of intervals where disruptive behaviour was observed at Time 2, and after four weeks, at Time 3.

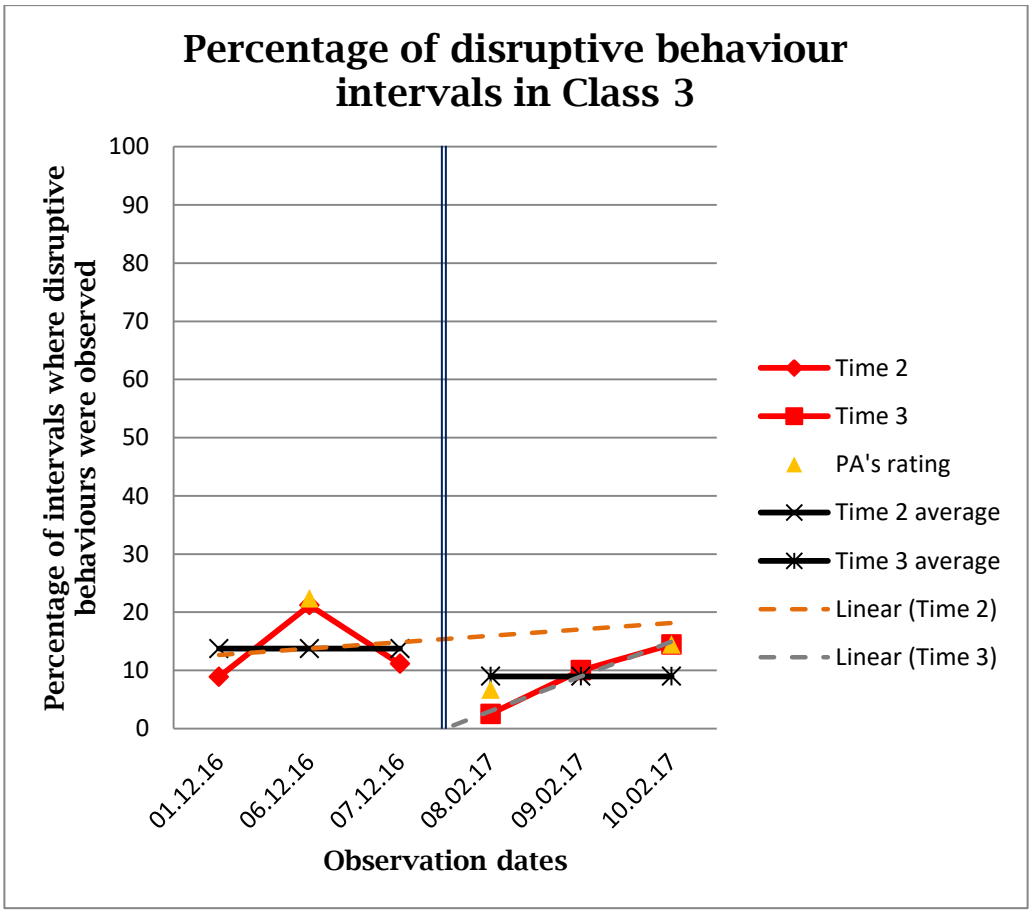


Figure 4.19: A graph to show the percentage of disruptive intervals observed in Class 3, at Time 2 and Time 3.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained disruptive behaviour at Time 2 was 13.73%. The mean percentage at Time 3 was slightly decreased to 8.98%. Mean change = 4.75%.
Trend	The Time 2 data showed an upward trend with a slope of 1.11. The Time 3 data showed a steeper upward trend, with a slope of 5.97. Slope change = 4.86.
Variability	Scores at Time 2 (range = 12.29% and standard deviation = 6.55%) were of a similar variability as scores at Time 3 (range = 11.94% and standard deviation = 6.03%).
Overlapping data points	2 data point from Time 3 overlapped with Time 2 data points. Percentage of overlapping data points = 66.66%

Table 4.13: A graph to show the percentage of disruptive intervals observed in Class 3, at Time 2 and Time 3.

The findings suggest that there was no real change in disruptive behaviour between Time 2 and Time 3, in Class 3. The change in level between the time points was small and both showed increasing trends. Furthermore, there was a large degree of overlap between the data points.

4.4.1.3. Class 4 (no change)

4.4.1.3.1. Implementation of the approaches

From observations and conducting a fidelity check, there was no evidence to suggest that there had been diffusion of treatment. The teacher reported being unaware of what the other classes were implementing and the other teachers reported that they had not shared with anybody else, what they were doing in their classrooms for this research.

4.4.1.3.2. Off-task behaviour

Figure 4.20 below shows the percentage of intervals where off-task behaviour was observed at Time 2, and after four weeks, at Time 3.

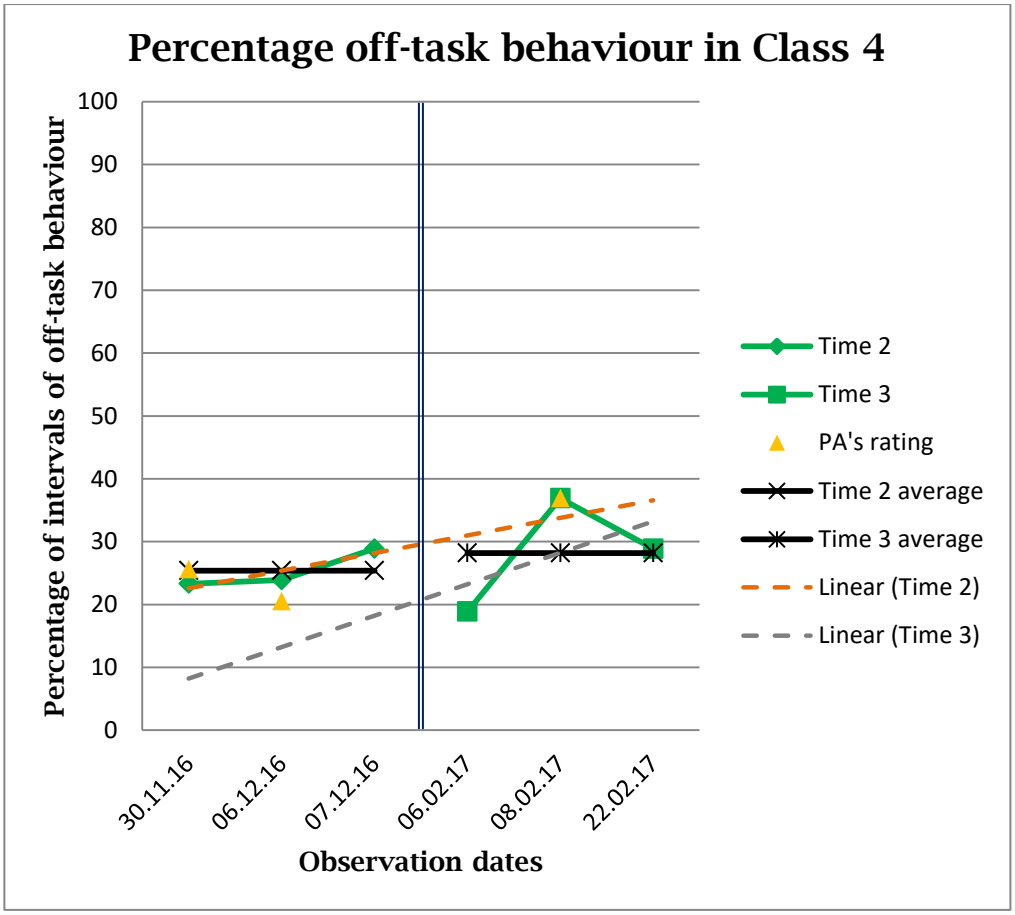


Figure 4.20: A graph to show the percentage of off-task intervals observed in Class 4, at Time 2 and Time 3.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained off-task behaviour at Time 2 was 25.4%. The mean percentage at Time 3 increased slightly to 28.2%. Mean change = 2.8%.
Trend	The Time 2 data showed an upward trend with a slope of 2.8. The Time 3 data showed a steeper upward trend, with a slope of 5.0. Slope change = 2.2.
Variability	Scores at Time 2 were much less variable, with a range of 5.6% and standard deviation of 3.07%. At Time 3, there was much more variability, with a range of 18.03%% and standard deviation of 9.03%.
Overlapping data points	1 data point from Time 3 overlapped with Time 2 data points. However all Time 2 data overlapped with Time 3 data. Percentage of overlapping data points = 33.33%

Table 4.14: A table to outline the findings from the visual inspection, on Class 4's graphed off-task behaviour data at Time 2 and Time 3.

The findings suggest that there was no change in off-task behaviour between Time 2 and Time 3 in Class 4. The difference in level between the time points was slight and both showed an increasing trend. Furthermore, data at Time 3 was highly variable and although only one Time 3 data point overlapped with the Time 2 data points, all of the Time 2 data points overlapped with the Time 3 ones.

4.4.1.3.3. Disruptive behaviour

Figure 4.21 shows the percentage of intervals where disruptive behaviour was observed at Time 2, and at Time 3, after four weeks.

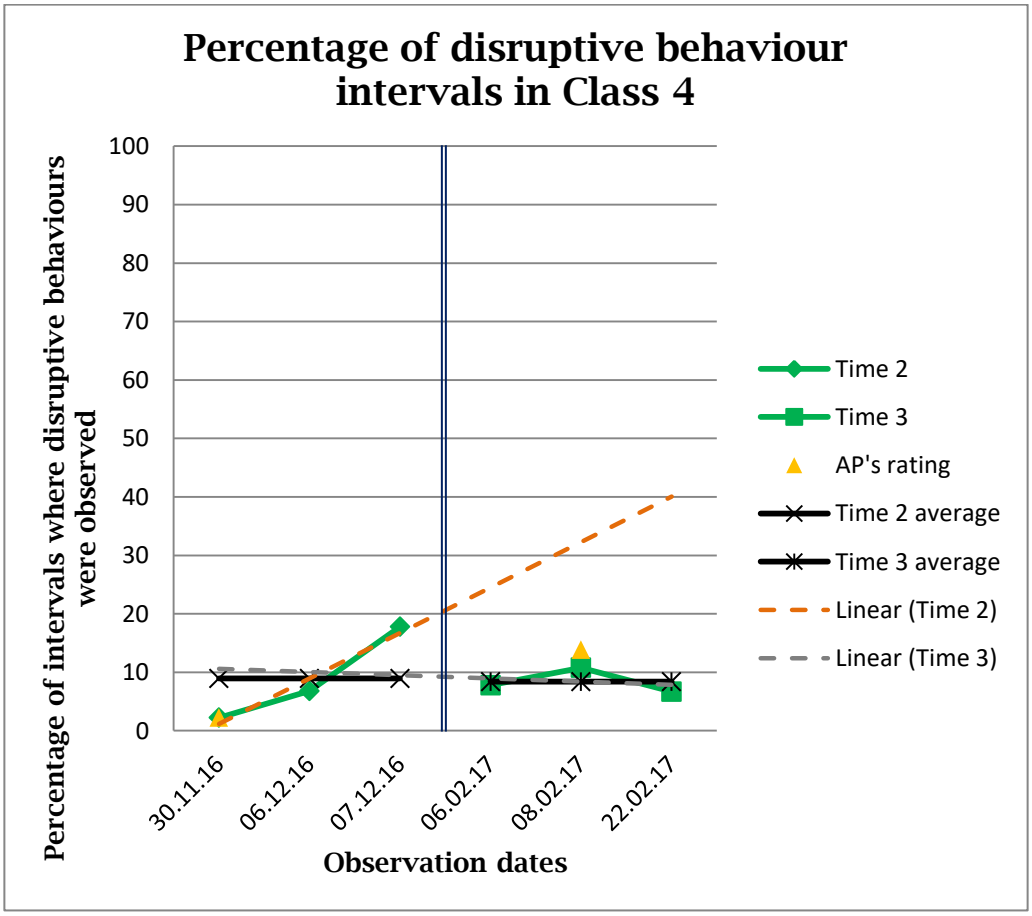


Figure 4.21: A graph to show the percentage of disruptive intervals observed in Class 4, at Time 2 and Time 3.

Visual Inspection	Findings
Level	The mean percentage of intervals which contained disruptive behaviour at Time 2 was 8.93%. The mean percentage at Time 3 was slightly decreased to 8.4%. Mean change = 0.53%.
Trend	The Time 2 data showed a steep upward trend with a slope of 7.78. The Time 3 data showed slight downward trend with a slope of -0.54. Slope change = 8.32.
Variability	Scores at Time 2 (range = 15.56% and standard deviation = 8.0%) were slightly more variable than scores at Time 3 (range = 4.07% and standard deviation = 2.11%).
Overlapping data points	All data point from Time 3 overlapped with Time 2 data points. Percentage of overlapping data points = 100%

Table 4.15: A graph to show the percentage of disruptive intervals observed in Class 4, at Time 2 and Time 3.

The findings suggest that there was no change in disruptive behaviour between Time 2 and Time 3, in Class 4. There was no real change in level between the data points and all of the Time 3 data points overlapped with those at Time 2. Also, although the trend changed from a steep increasing trend to a slight decreasing trend, the data at Time 2 was highly variable, which may suggest an inaccurate trend.

Figure 4.22 and Figure 4.23 combine the percentage of off-task and disruptive interval data respectively, from all three classes onto one graph, to aid visual comparisons between the groups.

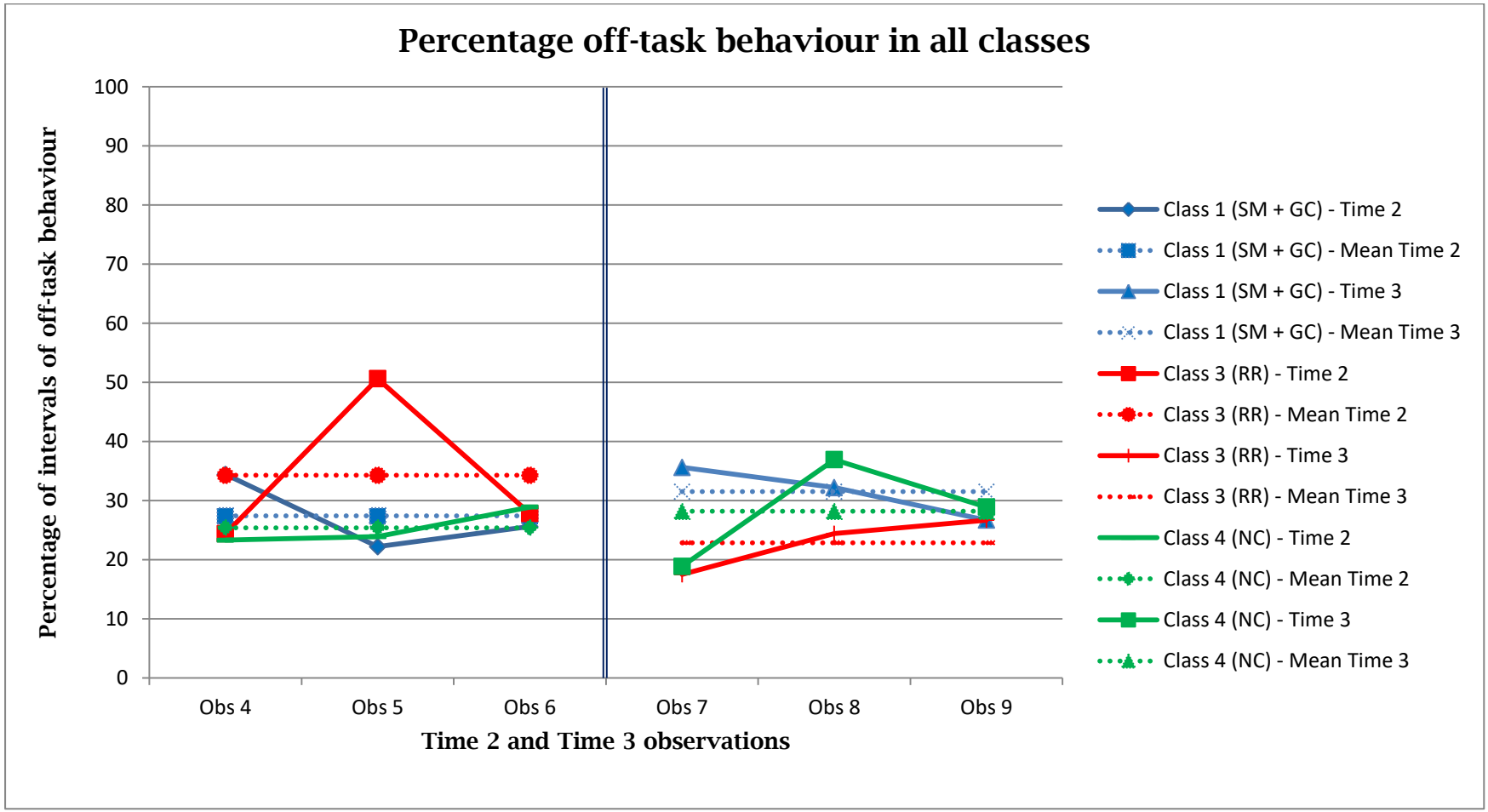


Figure 4.22: A graph to show the percentage of off-task intervals observed in classes 1, 3 and 4 at Time 2 and Time 3.

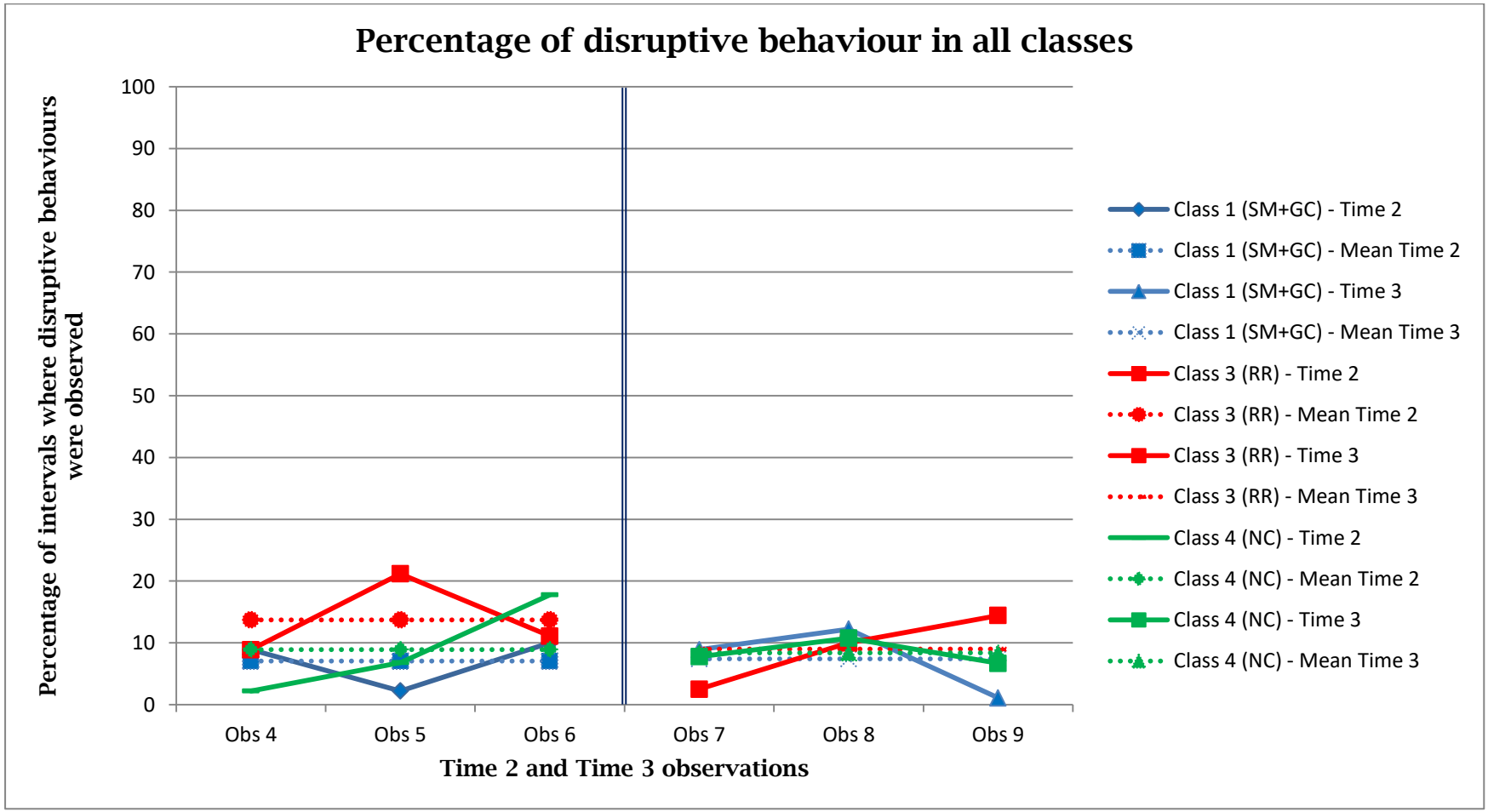


Figure 4.23: A graph to show the percentage of disruptive intervals observed in classes 1, 3 and 4 at Time 2 and Time 3.

Figure 4.22 illustrates that both Class 1 (SM+IGC) and Class 4 (NC) showed a slight increase in mean off-task behaviour at Time 3, suggesting that adding interdependent group contingency to an already existing self-management approach did not lead to further improvements on off-task behaviour. Class 3, however, did show a reduction in off-task behaviour which may suggest that being reminded of the rules daily, leads to improved engagement over time.

Figure 4.23 illustrates that after the combined approach had been implemented for four weeks, none of the groups differed from each other in their mean percentage of disruptive behaviour. Overall, the results suggest that adding interdependent group contingency to an already operating self-management intervention leads to no additional improvements in off-task and disruptive behaviour.

4.4.2. Strengths and Difficulties Questionnaire data

4.4.2.1. Results on the behavioural problems subscale

Figure 4.24 illustrates how Classes 1, 3 and 4 scored on behavioural problems at Time 2, before the approaches were combined in Class 1, and at Time 3, after four school weeks.

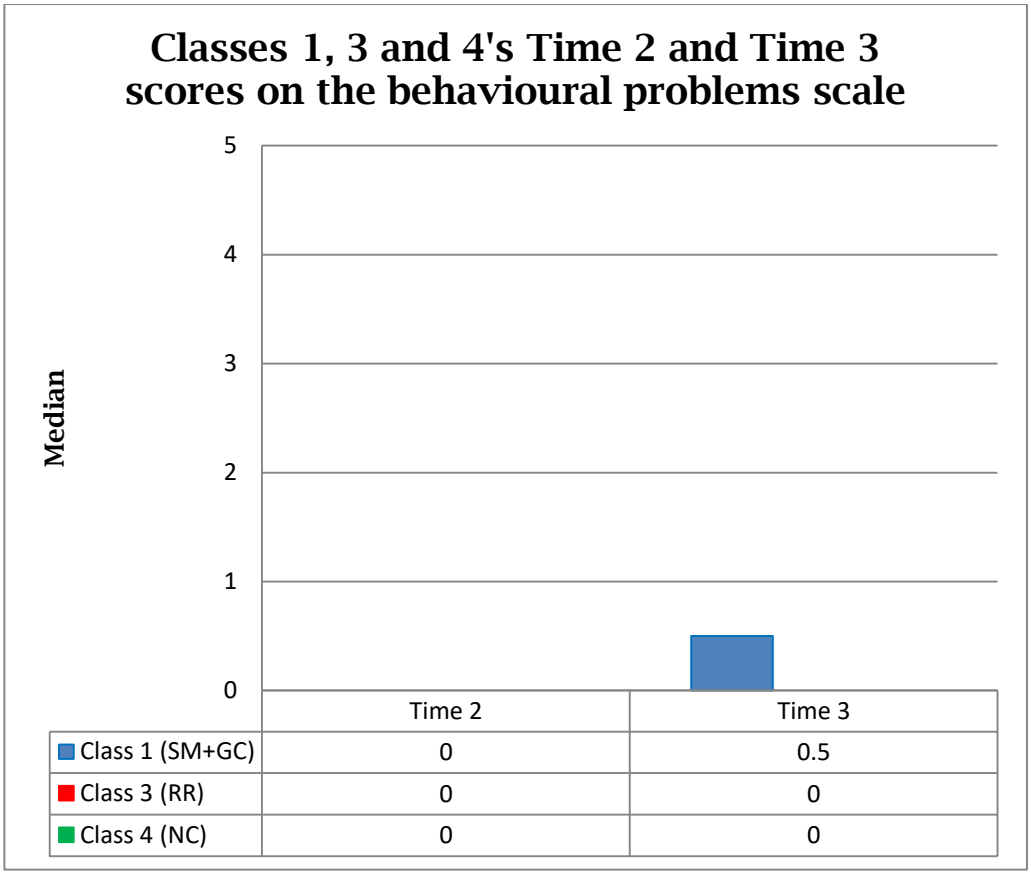


Figure 4.24: A graph to show class 1, 3 and 4's median scores on the behavioural problems subscale of the SDQ, at Time 2 and Time 3.

Kruskal-Wallis tests revealed that there was no significant difference between Classes 1, 3 and 4, on behavioural problems at Time 2 ($x^2(2)=2.774$, $p=.250$), or at Time 3 ($x^2(2)=6.817$, $p=.033$). The Wilcoxon Signed Rank Test highlighted that there was no significant difference in scores on behavioural problems between Time 2 and Time 3 in Class 1 ($z=-.237$, $p=.813$, $r=.03$), suggesting that implementing self-management with interdependent group contingency had no significant impact on behavioural problems. There was also no significant difference between Time 2 and Time 3 on behavioural problems for Class 3 (RR) ($z=-1.558$, $p=.119$, $r=.21$), or Class 4 (NC) ($z=-.530$, $p=.596$, $r=.07$). These effect sizes further indicate that there was no effect except in Class 3 (RR), where there was a small but non-significant change in behavioural problems.

4.4.2.2. Results on the hyperactivity and concentration problems subscale

Figure 4.25 illustrates how Classes 1, 3 and 4 scored on hyperactivity and concentration problems at Time 2, before the approaches were combined in Class 1, and at Time 3, after four school weeks.

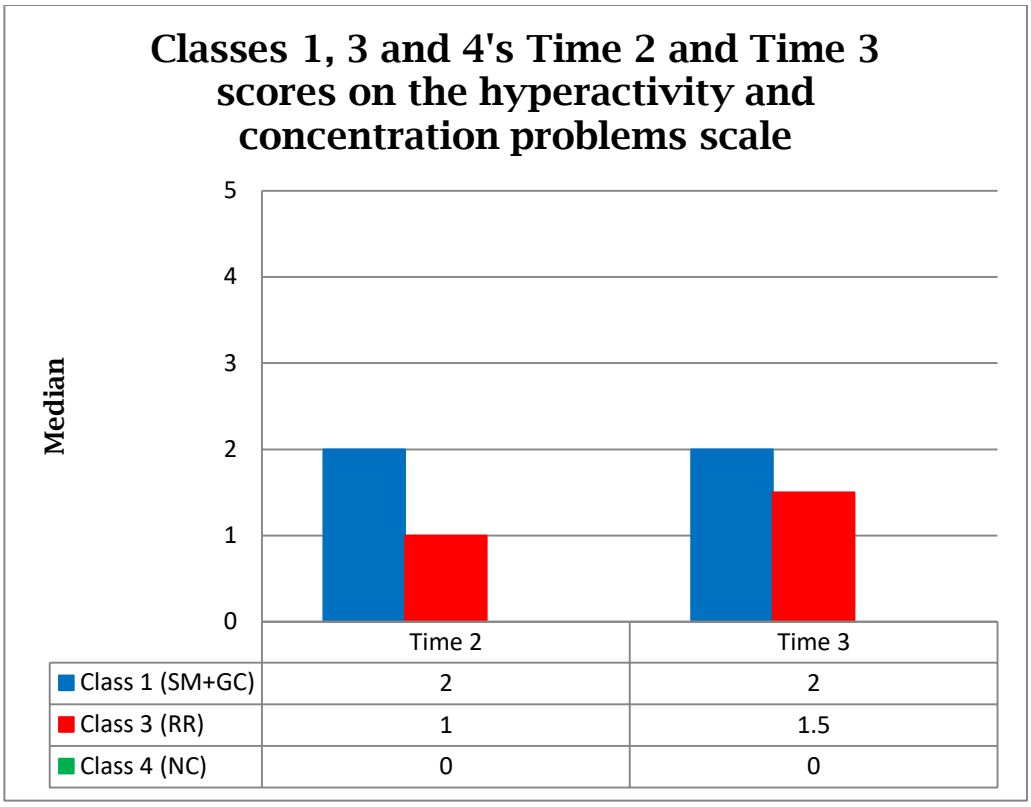


Figure 4.25: A graph to show class 1, 3 and 4's median scores on the hyperactivity and concentration problems subscale of the SDQ, at Time 2 and Time 3.

Kruskal-Wallis tests revealed that there were significant differences between Classes 1, 3 and 4, on hyperactivity and concentration problems at Time 2 ($x^2(2)=12.829$, $p=.002$), and at Time 3 ($x^2(2)=20.004$, $p=.000$). These findings therefore suggest that the groups were not equivalent at pre-test (Time 2) for this research question. As such, within-group comparisons have been conducted to highlight changes in each class between Time 2 and Time 3, with a calculation of effect size indicating the size of that change. The

Wilcoxon Signed Rank Test highlighted that there was no significant difference in scores on hyperactivity and concentration problems between Time 2 and Time 3 in Class 1 (z=-.211, p=.833, r=.03), suggesting that implementing self-management with interdependent group contingency had no significant impact on hyperactivity and concentration problems. There was also no significant difference between Time 2 and Time 3 on hyperactivity and concentration problems for Class 3 (RR) (z=-1.375, p=.169, r=.19), or Class 4 (NC) (z=-.681, p=.496, r=.09). These effect sizes further indicate that there was no effect except in Class 3 (RR), where there was a small but non-significant increase in hyperactivity and concentration problems.

4.4.3. **Overall summary of results for Phase B**

Overall, the findings from the observation data suggest that adding interdependent group contingency to an already operating self-management approach has no further impact on levels of off-task or disruptive behaviour. Furthermore, findings from the the SDQ suggest that implementing this combined approach also has no impact on general behaviour in terms of behavioural problems and hyperactivity and concentration problems, compared to implementing only self-management.

The next section discusses these findings in relation to the literature.

5. Discussion

5.1. Introduction

The previous chapter analysed and presented the results of this study in relation to the research questions (see Section 2.6). The current chapter begins by presenting the key findings for each research question in light of key literature and considers alternative interpretations. A discussion of the study's limitations is then presented before the implications of the findings for schools and EPs are outlined. Finally, avenues for future research are identified.

5.2. Summary of findings

5.2.1. Aims of Phase A

This study aimed to explore whether whole-class self-management or interdependent group contingency was most effective for reducing off-task and disruptive behaviours in a UK population.

The research also sought to investigate whether generalised improvements in behaviour were made, following these approaches, given that claims have been made that self-management and interdependent group contingency lead to maintained and generalised improvements in behaviour (Ennis et al., 2016; Freeman & Dexter-Mazza, 2004; Mitchem et al., 2001; Mooney et al., 2005).

5.2.2. Research question 1

Research question 1 asked: *What impact does whole-class self-management have on off-task and disruptive behaviours, as well as on the general behaviour of the whole class, as compared to control groups?*

5.2.2.1. Hypotheses

The hypotheses for research question 1 were:

i. *Pupils in the self-management group will show a decrease in off-task behaviour between Time 1 and Time 2, over and above those in the waitlist control conditions.*
ii. *Pupils in the self-management group will show a decrease in disruptive behaviour between Time 1 and Time 2, over and above those in the waitlist control conditions.*
iii. *Pupils in the self-management group will show a decrease in behavioural problems, as measured by the SDQ, between Time 1 and Time 2, over and above those in the waitlist control conditions.*
iv. *Pupils in the self-management group will show a decrease in hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2, over and above those in the waitlist control conditions.*

The null hypotheses were:

i. *Pupils in the self-management group will show no change in off-task behaviour between Time 1 and Time 2.*
ii. *Pupils in the self-management group will show no change in disruptive behaviour between Time 1 and Time 2.*
iii. *Pupils in the self-management group will show no change in behavioural problems, as measured by the SDQ, between Time 1 and Time 2.*
iv. *Pupils in the self-management group will show no change in hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2.*

5.2.2.2. Key findings and links to literature

The findings from this research suggest that self-management led to a reduction in off-task behaviours in the target lessons, which supports the findings of Mitchem et al., (2001). Both waitlist control groups showed no improvement at Time 2 compared to Time 1. This suggests

that it was insufficient to simply remind the class of the rules each day for four weeks. No difference in disruptive behaviour was observed as a result of implementing self-management, which contrasts the findings of Hoff and Ervin (2013), Mitchem and Young (2001), Niesyn (2009), and Bruhn et al., (2015) who all reported reductions in disruption. The questionnaire data suggests that self-management led to no significant improvements in general behavioural and hyperactivity and concentration problems, despite claims made that self-management *does* lead to maintained and generalised improvements in behaviour in some previous US-based research (Freeman & Dexter-Mazza, 2004; Mitchem et al., 2001; Mooney et al., 2005).

5.2.2.3. Alternative interpretations and limitations of this research question

These findings must be interpreted with caution, however, as they may also be explained by threats to internal validity that were present in the research. These factors are briefly outlined below with further discussion within the sign-posted sections.

- Off-task behaviour may have reduced in the self-management class due to a 'regression towards the mean' phenomenon rather than due to the implementation of the approach (see Section 5.4.1).
- The approach was not implemented as often as expected and this may explain why disruption, behavioural problems and hyperactivity and concentration problems did not decrease significantly (see Section 5.4.6).
- Time 1 scores for behavioural problems were close to floor levels, therefore it may not have been possible to detect improvements (see Section 5.4.3.2).
- The SDQ may not have been a sensitive enough measure when analysed at the whole class level, to detect small changes in generalised behaviour.
- Non-significant results may have been obtained for the SDQ data due to low statistical power of the tests (see Section 5.4.4).

5.2.2.4. Conclusion

The results indicate that null hypothesis *i* can be rejected as the class showed improvements in off-task behaviour compared to the control groups, however, null hypotheses *ii, iii* and *iv* must be accepted on this occasion, for these participants. Given the limitations of the research discussed above, however, these conclusions are tentative and further research exploring this question with a larger sample size, would be welcomed.

5.2.3. **Research question 2**

Research question 2 asked: *What impact does interdependent group contingency have on off-task and disruptive behaviours, as well as on the general behaviour of the whole class, as compared to control groups?*

5.2.3.1. Hypotheses

The hypotheses for research question 2 were:

i. *Pupils in the interdependent group contingency group will show a decrease in off-task behaviour between Time 1 and Time 2, over and above those in the waitlist control conditions.*
ii. *Pupils in the interdependent group contingency group will show a decrease in disruptive behaviour between Time 1 and Time 2, over and above those in the waitlist control conditions.*
iii. *Pupils in the interdependent group contingency group will show a decrease in behavioural problems, as measured by the SDQ, between Time 1 and Time 2, over and above those in the waitlist control conditions.*
iv. *Pupils in the interdependent group contingency group will show a decrease in hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2, over and above those in the waitlist control conditions.*

The null hypotheses were:

i. *Pupils in the interdependent group contingency group will show no change in off-task behaviour between Time 1 and Time 2.*
ii. *Pupils in the self-management group will show no change in disruptive behaviour between Time 1 and Time 2.*
iii. *Pupils in the interdependent group contingency group will show no change in behavioural problems, as measured by the SDQ, between Time 1 and Time 2.*
iv. *Pupils in the interdependent group contingency will show no change in hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2.*

5.2.3.2. Key findings and links to literature

The findings from this research suggest that interdependent group contingency leads to a reduction in off-task and disruptive behaviours, which supports a wealth of prior research (Christ & Christ, 2006; Ennis et al., 2016; Hansen & Lignugaris-Kraft, 2005; Hartman & Gresham, 2016; Kelshaw-Levering et al., 2000; Ling & Barnett, 2013; Ling et al., 2011; McKissick et al., 2010; Murphy et al., 2007; Theodore et al., 2004; Williamson et al., 2009). The reduction in off-task behaviours was smaller than the reduction in disruptive behaviours. This is in comparison to both control classes who showed no reductions in off-task or disruptive behaviour.

In contrast, the data from the questionnaires suggested that following implementing interdependent group contingency, the class experienced significant increases in hyperactivity and concentration problems, and noticeable but statistically non-significant increases in behavioural problems, relative to controls. This suggests that interdependent group contingency led to greater general difficulties with hyperactivity and concentration, and possible deterioration in behaviour as well. This was unexpected as it contrasts previous research in which maintained and generalised improvements in behaviour were reported (Ennis et al., 2016).

5.2.3.3. Alternative interpretations and limitations of the research

Again, these findings must be interpreted with caution as they may also be explained by threats to internal validity that were present in this research. These factors are briefly outlined below with further discussion within the sign-posted sections.

- The improvements in off-task and disruptive behaviour may have occurred due to maturation, given that the teacher's absence meant Time 2 data collection occurred seven weeks after the other classes' data was collected (see Section 5.4.8).
- Time 2 data in the control classes was collected in the two weeks before the Christmas holidays, during unstructured activities. This may have negatively impacted the control classes' data in comparison to interdependent group contingency (see Section 5.4.8). This may have inaccurately suggested that the interdependent group contingency class showed more improvement than the control groups.
- The teacher SDQ may have inaccurately rated the pupils' general hyperactivity and concentration, and behavioural problems as worse, possibly due to teacher bias or inaccurate rating through hasty completion of the questionnaires (see Section 5.4.3.2). Furthermore, this data did not triangulate with observation data, which suggested improvements in Class 2 at Time 2.
- The class may have experienced increased hyperactivity and concentration problems, and behavioural problems as a result of the change in routine and teaching/management style experienced from the teacher's long absence (see Section 5.4.8).
- The class may have experienced no significant change in behavioural problems as a result of the Time 1 scores being close to floor levels, meaning there was limited capacity to measure improvement.

5.2.3.4. Conclusion

The results indicate that null hypothesis *i* and *ii* can be rejected as the class showed improvements in off-task and disruptive behaviour compared to the control groups. The results indicate that null hypotheses *iv* must also be rejected as it appears that the approach may have had a negative impact upon general hyperactivity and concentration problems. Null hypothesis *iii,* however, can be accepted as there was no statistically significant difference, compared to control groups, on behavioural problems following the implementation of the approach. Given the limitations of the research discussed above however, these conclusions are tentative and further research exploring this question with a larger sample size, would be welcomed.

5.2.4. **Research question 3**

Research question 3 asked: *Which approach (self-management or interdependent group contingency) is most effective in reducing off-task and disruptive behaviours, and in improving general behaviour?*

5.2.4.1. Hypotheses

The hypotheses for research question 3 were:

i. *Pupils in the self-management condition will show equivalent reductions in off-task behaviour to pupils in the interdependent group contingency condition, between Time 1 and Time 2.*
ii. *Pupils in the self-management condition will show equivalent reductions in disruptive behaviour to pupils in the interdependent group contingency condition, between Time 1 and Time 2.*
iii. *Pupils in the self-management condition will show a greater decrease in behavioural problems, as measured by the SDQ, between Time 1 and Time 2, compared to pupils in the interdependent group contingency condition.*
iv. *Pupils in the self-management condition will show a greater decrease in hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2, compared to pupils in the interdependent group contingency condition.*

The null hypotheses were:

i. *Pupils in both the self-management and the interdependent group contingency conditions will show no change in off-task behaviour, between Time 1 and Time 2.*
ii. *Pupils in both the self-management and the interdependent group contingency conditions will show no change in disruptive behaviour, between Time 1 and Time 2.*
iii. *Pupils in both the self-management and the interdependent group contingency conditions will show no change in behavioural problems, as measured by the SDQ, between Time 1 and Time 2.*
iv. *Pupils in both the self-management and the interdependent group contingency conditions will show no change in hyperactivity and concentration problems, as measured by the SDQ, between Time 1 and Time 2.*

5.2.4.2. Key findings and links to literature

The findings from this research suggest that self-management leads to greater improvements for off-task behaviour than interdependent group contingency, which supports the findings from Glynn et al., (1973); however, it appears that interdependent group contingency is more effective for reducing disruptive behaviour, even though this reduction in disruptive behaviour was small. These findings contrast Hoff and Ervin (2013), who concluded that they were unable to establish which approach had the greatest impact on behaviour. The findings also indicate that both self-management and interdependent group contingency led to no significant improvements in general behavioural problems; behaviour appeared to deteriorate significantly between Time 1 and Time 2 in Class 2, although the Time 2 scores in both classes did not significantly differ. With hyperactivity and concentration problems, self-management showed no change, whereas interdependent group contingency showed significantly increased scores.

These findings suggest that self-management is preferable for off-task behaviour and interdependent group contingency is preferable for disruptive behaviour. Neither approach, however, appears to be

effective in supporting generalised improvements in behaviour as claimed in the literature (Ennis et al., 2016; Freeman & Dexter-Mazza, 2004; Mitchem et al., 2001; Mooney et al., 2005). Overall it appears that the limited research conducted to compare whole-class self-management with interdependent group contingency yields mixed results and would benefit from further exploration.

5.2.4.3. Alternative interpretations and limitations of the research

Previous sections (Sections 5.2.2.3 and 5.2.3.3) outlined some threats to the validity of the results, which may have impacted on the overall findings for this research question. Additional alternative interpretations and limitations are outlined here:

- The greater disruptive behaviour improvements observed in the interdependent group contingency class may have been due to maturation with this class' data being collected seven weeks later, or due to the approach being implemented for an additional week compared to self-management (see Section 5.4.8).
- Time 2 data in the self-management class was collected in the two weeks before the Christmas holidays, when unstructured activities were taking place in school (see Section 5.4.8). This may have inaccurately suggested that the interdependent group contingency class showed more improvement on disruptive behaviour than the self-management class.
- Non-significant results may have been obtained for the SDQ data due to low statistical power of the tests (see Section 5.4.4).

5.2.4.4. Conclusion

The results indicate that null hypothesis *i* and *ii* can be rejected as interdependent group contingency showed improvement on both off-task and disruptive behaviour, and self-management appeared to lead to greater improvements on off-task behaviour. Null hypotheses *iv* must also be rejected as interdependent group contingency showed significantly higher scores at Time 2 compared to Time 1 on

hyperactivity and concentration problems. Null hypothesis *iii* can only be partially accepted, as interdependent group contingency showed significant increases in behavioural problems between Time 1 and Time 2, whereas self-management showed no significant change.

5.2.5. Phase A: Further links to the literature

As previously reviewed psychological theory and research state, the improvements in off-task and disruptive behaviour observed in this research may have occurred, in general terms, as a result of the pupils being motivated to follow the rules (Lewin, 1935), through these behaviours being positively reinforced with rewards (Gleitman et al., 2004; B. F. Skinner, 1953), through developing social responsibility among group members (D W Johnson, 2003), and through the goals of the teacher and those of the pupil aligning (Deutsch, 1949). Perhaps greater improvements were not observed due to incompatibilities between the teacher and pupil goals (D W Johnson & Johnson, 2006), due to the rewards not being intrinsically motivating (Deci & Ryan, 1985; Ryan & Deci, 2000), or due to disruptive behaviours possibly being reinforced through peer attention (Altman & Linton, 1971; Northup et al., 1995).

It is important to look more specifically at potential reasons for why pupils' behaviour in class appeared to respond differently to the two different approaches. The findings suggest that self-management led to greater improvements in engagement than interdependent group contingency, but that only interdependent group contingency was effective for reducing disruptive behaviour. It may be that off-task behaviour and disruptive behaviour have different underlying causes and as a result, respond differently to the different approaches. The fact that only interdependent group contingency reduced disruptive behaviours might suggest that disruptive behaviour in this population occurred due to being positively reinforced in some way, more so than following the rules (B. F. Skinner, 1953). As such, the reward for following the rules and the developed social responsibility to satisfy the interdependent interests of friends (Deutsch, 1949; D W Johnson & Johnson, 2005), may have reduced disruption. In further support of

this hypothesis, self-management alone placed no wider responsibility on an individual to satisfy the collective desires of the group and there was no reward incentive for following the class rules, to make it a more desirable option. Freeman and Dexter-Mazza (2004) argued that quality feedback is essential to developing self-regulation skills. In this study, the pupils provided feedback to each other rather than the teacher providing this, and this may have compromised the quality of that feedback. This may explain why disruption did not reduce with self-management.

Off-task behaviour may not occur entirely due to being positively reinforced and therefore may not be reduced through simply increasing motivation to remain *on*-task through extrinsic rewards. This supports Appleton et al., (2008) who stated that motivation is necessary, but is alone insufficient to increase engagement. If off-task behaviour is caused by difficulties with self-regulation as hypothesised by self-regulation theory (Bandura, 1991; Zimmerman, 2000), one would expect reductions in off-task behaviour to be reduced more so following self-management than interdependent group contingency, as was the case in this research. Self-management prompts pupils to think about what they are doing and what they should be doing in pursuit of their 'goal' in that lesson (Bandura, 1991; Zimmerman, 2000), thereby returning their attention to the task.

These suggested reasons for the difference in findings between self-management and interdependent group contingency are avenues to explore with future research; this study did not seek to explore the mechanisms or processes by which these approaches impacted classroom engagement and disruption. As such, the possible explanations presented here are speculative.

5.2.6. **Aims of Phase B**

Phase B aimed to identify whether combining the two approaches led to any further reductions in off-task and disruptive behaviour as well as improvements in general behaviour, which might support claims that the approaches are most effective when delivered together (Freeman & Dexter-Mazza, 2004; Maggin et al., 2012; Reiber &

McLaughlin, 2004). The study sought to build on the research by Denune et al., (2015) and Hoff and Ervin (2013) who concluded that adding self-management led to no further improvements in behaviour after interdependent group contingency had been implemented. Generalised changes in behaviour were measured using the SDQ.

5.2.7. **Research question 4**

Research question 4 asked: *Is there an added benefit to combining self-management with interdependent group contingency, with regard to off-task and disruptive behaviours, as well as general behaviour?*

5.2.7.1. Hypotheses

The hypotheses for research question 4 were:

i. *Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show a reduction in off-task behaviour between Time 2 and Time 3, which is greater than those observed in the control conditions.*
ii. *Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show a reduction in disruptive behaviour between Time 2 and Time 3, which is greater than those observed in the control conditions.*
iii. *Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show a reduction in behavioural problems between Time 2 and Time 3, which is greater than those observed in the control conditions.*
iv. *Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show a reduction in hyperactivity and concentration problems between Time 2 and Time 3, which is greater than those observed in the control conditions.*

The null hypotheses were:

i. *Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show no change in off-task behaviour between Time 2 and Time 3.*
ii. *Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show no change in disruptive behaviour between Time 2 and Time 3.*
iii. *Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show no change in general behavioural problems, as measured by the SDQ, between Time 2 and Time 3.*
iv. *Pupils who receive the combined approach (interdependent group contingency alongside self-management), will show no change in hyperactivity and concentration problems, as measured by the SDQ, between Time 2 and Time 3.*

5.2.7.2. Key findings and links to literature

Findings suggest that adding interdependent group contingency to an already operating self-management approach led to no difference in off-task behaviour or disruptive behaviour in the target lessons. This contrasts with findings from Ardoin and Martens (2004), and Graham-Day et al., (2010) investigating the impact on individual pupils, and with Chafouleas et al., (2012), who reported that adding an interdependent group contingency led to greater reductions in whole-class off-task behaviour. The findings also appear to counter claims by Reiber and McLaughlin (2004) who stated that such improvements may occur due to increased motivation from receiving a reward.

The findings suggest that combining these two approaches leads to no improvement in general behavioural problems or hyperactivity and concentration problems. It therefore appears that there is no additional benefit to adding interdependent group contingency to self-management and that combining the approaches is no more effective than implementing self-management alone.

Unfortunately, due to unforeseen circumstances (see Section 3.5.2), it was not possible to collect Class 2's Time 3 data on the impact of adding self-management to interdependent group contingency (controlling for ordering effects). Prior research, however, suggests that adding self-management to interdependent group contingency may have had no further impact on off-task and disruptive behaviour (Denune et al., 2015; Hoff & Ervin, 2013).

5.2.7.3. Alternative interpretations

These findings may, to some degree, be explained by threats to internal validity that were present in this research. These factors are briefly outlined below with further discussion within the sign-posted sections.

- Adding interdependent group contingency to an already operating self-management approach may have led to no further improvements in engagement and disruptive behaviour or general behavioural and hyperactivity and concentration because this combined approach was not implemented daily as intended (see Section 5.4.6).
- Unlike with Time 1 and Time 2 data the majority of Time 3 data was collected in the second lesson of the afternoon. This may have added an additional variable that affected the results and comparability of the data (see Section 5.4.1).
- The final data point for Class 1 (SM+IGC) and Class 4 (NC) was collected after the one week half-term break, which may have introduced a confounding variable.
- Behavioural problems as measured by the SDQ were already at floor levels at Time 2. As such, the fact that no significant difference could be observed on this measure at Time 3 may be because there was no capacity to measure such improvement (see Section 5.4.3.2).
- Non-significant results may have been obtained for the SDQ data due to low statistical power of the tests (see Section 5.4.4).

5.2.7.4. Conclusion

The results indicate that null hypotheses *i, ii, iii* and *iv* cannot be rejected as the data in this study does not indicate that combining self-management with interdependent group contingency leads to further improvements in off-task or disruptive behaviour in the target lessons or on general behaviour outside of these lessons.

Although all of the null hypotheses cannot be rejected for this research question given the findings presented, it is important to note that the data for this research question was obtained from only one class, and in that one class the approach was not implemented as often as was expected. As such, although the data suggests that there is no added benefit to combining the two approaches, such a conclusion is tentatively made and the limited data is acknowledged as being insufficient to answer the research question with confidence. What this study *can* conclude is that implementing the approach three times over four weeks is unlikely to be sufficient for improvements to be observed. Further research which aims to answer this research question with a larger sample and daily implementation would be welcomed.

5.3. A wider focus on factors affecting engagement and disruption

Self-management and interdependent group contingency sought to improve engagement and reduce disruption through focusing on and reinforcing certain desired behaviours (B. F. Skinner, 1953), and through teaching pupils to monitor and change, as necessary, their own behaviour (Freeman & Dexter-Mazza, 2004; King-Sears, 2008; Niesyn, 2009; Rooney et al., 1984; Traxson, 1994). As such, the focus was entirely on behaviour regulation. However, many have argued for the need to take a broader, more holistic view of behaviour in the context of its environment, such as examining school climate and communication (Reschly & Christenson, 2012).

Although psychological theory and research does suggest that classroom behaviour can be improved through increasing motivation with rewards (Altman & Linton, 1971; Bednar et al., 1970; Burchard & Tyler, 1964; McAllister et al., 1968; Osborne, 1968), through increasing social interdependence (Deutsch, 1949; D W Johnson & Johnson, 2005), and through equipping pupils with the skills to reflect on their own behaviours (Bandura, 1991; Zimmerman, 2000), there is a large body of research that highlights the central role that the classroom context plays in affecting engagement and disruptive behaviour (Fredricks et al., 2004; M. M. Mitchell & Bradshaw, 2013; Stronge et al., 2004), and this supports self-determination theory (Ryan & Deci, 2000). For instance, disengagement may have continued in these classes due to the quality of the teacher-pupil relationship and interaction (Evertson & Weinstein, 2006; Hajdukova et al., 2014; Marsh, 2012; O'Connor et al., 2011), the degree to which the teachers were autonomy-supportive (Jang et al., 2010; F. Mitchell et al., 2015; Shih, 2008; Standage et al., 2005; Stroet et al., 2013; Van den Berghe et al., 2016), the degree to which they gave positive feedback (Deci et al., 1999; Koka & Hagger, 2010), and the degree to which they supported pupils in feeling competent at a task (Garon-Carrier et al., 2015; Lavigne et al., 2007; W. Lee et al., 2014). If the task is presented in a way that undermines these needs, motivation may reduce (Reeve, 2012), and this may explain why disengagement was not reduced further.

Additional influencing factors include class size (Blatchford et al., 2011), peer problems (Buhs et al., 2006), classroom structures and task characteristics (Fredricks et al., 2004). Psychological theory and research suggests that children seek to have their need for autonomy, competence and relatedness met (Deci & Ryan, 1985; Reschly & Christenson, 2012; Ryan & Deci, 2000), and self-management aims to support pupils in being autonomous and competent rule-followers. Perhaps this focus is incorrect. Perhaps the need for autonomy, relatedness and competence should be met through how the learning is presented (as the literature suggests), rather than focusing on rule-following behaviour (as self-management aims to), to increase engagement.

It was not the objective of self-management and interdependent group contingency approaches to intervene with these wider key factors that can affect behaviour and this may therefore explain why disengagement and disruption were not reduced further. Self-management seeks to develop self-regulation skills and lacking the skills to regulate may arguably have been a cause for disengagement (Carter & Doyle, 2006; Lovitt & Curtiss, 1969; McLaughlin, 1976), but it is unlikely to have been the only cause.

Overall it appears that although interdependent group contingency and self-management intervene with important factors that affect engagement and behaviour such as motivation, reward and self-regulation skills, it is insufficient to focus only on these factors and doing so may limit potential increases in engagement and reductions in disruption.

5.4. Limitations of the research

The aim of experimental research is to establish a cause and effect relationship between the independent variable (IV, in this case, the whole class behaviour management approaches implemented), and the dependent variable (DV, in this case, off-task and disruptive behaviour) (Mertens, 2005; Robson, 2011). This causal link is strengthened when other factors are controlled for (Mertens, 2005; Robson, 2011), however, in real world research, despite best efforts, this is rarely possible (Robson, 2011). Often, extraneous variables introduce threats to the reliability and internal validity of the research (Robson, 2011). This means the conclusions drawn from the findings must be tentative and interpreted with caution (Mertens, 2005). This section outlines the limitations of the current study, which despite best efforts to avoid, may have impacted on the results.

5.4.1. Research design

The quasi-experimental design of the research meant that each of the four groups were taught by different teachers, each with their own individual teaching style, classroom management style and routines. This could not be controlled for and therefore may have impacted on

the pupils' observed and reported behaviour. Research highlights the impact that the teacher-pupil relationship can have on classroom behaviour (Evertson & Weinstein, 2006; Hajdukova et al., 2014; Marsh, 2012).

Furthermore, despite originally agreeing with teachers to implement the approaches and conduct the observations in the 13.30-14.30 slot, to control for time-of-day factors such as tiredness, it was often not possible for the teachers to implement the approach in that time period. On some occasions the approach was instead conducted in the 14.30-15.30 hour, which added additional variables that could have affected the results. This was problematic because the majority of Time 1 data was collected in the 13.30-14.30 slot. Time 2 data was collected during both the first and second hour, and due to Time 3 data being collected during 'book week' where each afternoon began with an assembly, the majority of Time 3 data was collected in the 14.30-15.30 slot. On reflection, it would have been beneficial to have consulted the class timetable in advance of the research and to have problem-solved with teachers any potential obstacles to daily implementation. The researcher may instead have suggested that the approaches be implemented in a morning lesson where the children are in ability classes, possibly taught by a different teacher, but have the form tutors fill out the SDQ to measure the children's general behaviour outside of the morning lesson. This may also have reduced the risk of individual teacher bias, as the questionnaires for each class would not all be completed by the same teacher.

The research design also meant that classes were observed in a variety of lessons undertaking different types of tasks. These different lessons and tasks types may have impacted on how the pupils differentially behaved during each observation, in relation to how their needs were being met in those lessons. This may account for some of the variance in the observed behaviours, separate to the classroom management approach being implemented. It would have been better to have collected more observation data at each time point, to control for task types and to give a more reliable indication of where the true level of engagement and disruption was for the classes. Alternatively, focusing

the approach in numeracy or literacy only would have allowed for daily implementation and for task types to be better controlled for.

The researcher could have asked the PAs to conduct observations in the other classes simultaneously, so that more observation data could be collected, however, there was insufficient time and resources to train them to the appropriate level, to ensure reliability of the data collected and to avoid observer error and bias. With the researcher collecting all data, the observer variable was controlled for. On the other hand, as the researcher was also the data collector, blind observations were not conducted and so even though attempts to control for this were made through IOA checks, it cannot be ruled out that there may have been some observer bias in the data collection. On reflection, the researcher could have planned for extensive training on the observation schedule to be delivered to the PAs from September, while parental consent was being sought. This would have meant that the PAs were appropriately trained to collect data without the researcher. This would have enabled more observations and more reliable data to be obtained.

In real world research it is often not possible to randomly allocate participants to conditions, and group differences at pre-test can therefore sometimes be observed. This was the case in this research. Despite choosing the same year group and school, the pupils did differ in their Time 1 scores for the observation data. This may indicate that the pupils and/or teachers in each group differed in characteristics. As such, the experimental class may have responded in the observed way due to their individual group characteristics rather than as a result of the approaches. This increases the threat of *differential selection* to the internal validity of the results. This difference at Time 1 may also have meant that Class 1 (SM) and Class 2 (IGC) improved not due to the approach, but because their scores were 'more extreme', leading to a regression towards the mean phenomenon. This is where through re-testing, extreme scores become less extreme (Imai, 2017). Also, having a higher initial score gave these classes a greater capacity to demonstrate improvement than the control classes. The improvements

observed in Class 1 and Class 2 at Time 2 became equivalent to the scores recorded for the control classes.

An additional limitation to this study was that maintenance was not measured. Due to time limitations, it was not possible to return to collect follow-up data.

A single-case experimental design or multiple baseline design may have been more effective in establishing a causal link, through being able to implement the approach in every class and stagger when the approach was implemented. This type of research would have been beyond what was possible to achieve in this small-scale study as observations would need to have been conducted at least three times a week in *every* class for at least eight weeks.

5.4.2. **Behaviour of the classes**

The observations conducted in the classrooms highlighted that rather than most of the pupils displaying off-task or disruptive behaviour, in reality, the majority of those behaviours was exhibited by a small number of individuals. As such, the behaviour of a few pupils may have skewed the classes' data. If it was possible to have scored all children at all times, this would have given a truer estimate of the classes' overall behaviour. Given that many of the children displayed on-task behaviour much of the time, it is unsurprising that the majority of individuals in each class scored low on the hyperactivity and concentration, and behavioural subscales of the SDQ. Also, it is likely to explain why off-task and disruptive behaviour scores were not higher at Time 1 than was recorded in this research. As such, the capacity to observe improvements of the whole class would have been greatly reduced, given that the many who exhibited on-task behaviour at Time 1 would not have been able to improve further. This could have led to the true impact of self-management and interdependent group contingency being underestimated. It would have been useful to have measured the impact of these approaches on only the pupils who exhibited the greatest behavioural challenges, to explore the efficacy of these approaches for that population. On reflection, a criterion for

choosing classes should have been that the majority of students must display off-task or disruptive behaviour most of the time.

5.4.3. Measures

5.4.3.1. Observation schedule

Although there exists no 'gold standard' observation tool for coding the behaviour of a whole class, the partial interval recording schedule was sensitive enough to detect subtle changes in behaviour. However, as this provided only an estimate of actual behaviour (Clemens et al., 2013; Rapp et al., 2008), it may have provided unreliable data. The scores for disruptive behaviour may also have been unreliable as it was not possible to reliably measure the number of instances of disruptive behaviour occurring in the class as a whole. As such, disruptive behaviour was only counted if the pupil being observed was being disruptive during that interval or if they were being disrupted by somebody else in that interval. As such, all other instances of disruption in the class were not detected. As the observations were conducted in a fixed order by table, if the pupil observed happened to be disruptive/disrupted, the chances of the next two or three intervals containing disruption were increased by virtue of those pupils all being on the same table and perhaps all disruptive/disrupted in that moment. If another table had been observed during those 2 minutes, those instances of disruption would not have been recorded. This may have led to an over- or under-estimation of disruptive behaviours. It may have been better to have had two observers; one who focused on off-task behaviour only and one who focused on disruptive behaviour, again following a fixed schedule but observing groups rather than individuals, to capture as much disruption as possible. Alternatively, the researcher could have alternated between observing off-task behaviour and disruptive behaviour, as described, by conducting one for half of the observation, before swapping.

One of the difficulties faced with observing off-task and disruptive behaviour in the lessons was trying to balance being an 'invisible' observer who has minimal impact on the naturalistic setting, with

obtaining a reliable and clear picture of whether a behaviour was task-related or not. The most difficult aspect of coding behaviour was trying to judge whether pupils were talking about their work (on-task) or about something off-topic (off-task), when it was not possible to move closer to them to listen in. Visual clues were used to make a best guess (see Section 3.5.8.2.2), however this may have led to incorrect coding, impacting the validity of the findings. It is likely that despite standing in one corner of the room, out of the direct eye line of most pupils, the pupils' behaviour may have changed due to there being an observer in the room. At times, the pupils would turn to look at the researcher during the observations. As such, the observations may not have captured the 'true' behaviour of the class.

5.4.3.2. Strengths and Difficulties Questionnaire

Despite the high levels of reliability and validity reported for the SDQ, it may not have been entirely appropriate to use the *behavioural problems* subscale for this research. This subscale consists of five items, two of which were not relevant to classroom engagement. These were:

- *"Often lies or cheats"*
- *"Steals from home, school or elsewhere"*

These had to be included, however, as it would have been inappropriate to only score some items within a subscale. The pupils in all classes scored close to floor levels on this subscale at Time 1, which limited the capacity to measure much improvement, meaning it may not have been the most suitable measure of generalised behaviour. The SDQ however was sensitive enough to detect changes between Time 1 and Time 2 on the *hyperactivity and concentration problems* subscale, which was useful for this research.

The SDQ was helpful in gaining an overall picture of a pupil's classroom behaviour from the adult who knows them best, however, being a *teacher-reported* measure, it was vulnerable to *participant error* or *bias;* that is the way the teachers responded to the questionnaire may have been affected by *when* they were filling them out. For

instance, as the Time 2 data was collected in three classes close to Christmas and the Time 3 data was collected in those same classes just before the February half term, teachers may have responded differently due to being tired or having larger workloads at that time. Had the teachers filled them out on a different day, the outcomes may have been different. This may especially have been the case in Class 2 (IGC), where the teacher had not completed 21 of the 29 questionnaires by the agreed date. They may have been completed in greater haste than the previously completed questionnaires, and perhaps without the same level of consideration. The teacher rating may have reflected their perception of the class on that day rather than in general. This may explain why Class 2 (IGC) scored significantly higher on hyperactivity and concentration problems at Time 2 compared to Time 1. This possibility is further strengthened by the fact that observation data for this class at Time 2 indicated improved behaviour.

It would have been useful to have collected additional, more objective data to triangulate with the teacher reported data. This may have been in the form of additional observations in non-target lessons, but would have required additional resources which were not practical given the scale of the present project.

This research measured engagement through observing the behaviour of the pupils in each class, however on reflection, it would have been useful to also have measured emotional and cognitive engagement (Fredricks et al., 2004), through pupil self-report measures, in order to get a fuller measure of engagement as a construct.

5.4.4. **Data analysis**

A further limitation of the research was the inability to use a parametric test to analyse the SDQ data. As there does not exist a non-parametric alternative to the ANOVA, multiple tests had to be conducted and so the alpha level was adjusted to p=.01 instead of p=.05, which may have increased the likelihood of making a Type 2 error (incorrectly accepting the null hypothesis), due to reducing the power of the non-parametric tests used. Small sample sizes in each

group is likely to have further reduced the power of the tests (Field, 2005; Pallant, 2007; Stevens, 2012).

Given that individual pupils were not randomly allocated to conditions in this research, it may be that the scores obtained for each pupil on the SDQ were not independent, thereby violating the assumption of independence of observation which is required when conducting statistical tests (W. E. Martin & Bridgmon, 2012). Independence may have been violated because the pupils in each group were more similar, having come from the same class (Shadish et al., 2002), than would have been the case if they had been randomly allocated to conditions (Shadish et al., 2002). Additionally, being in the same class may have led to the pupils influencing each other's behaviour, therefore affecting their individual SDQ scores. Having the same teacher and experiencing the same classroom environment may also have led to related rather than independent scores on the SDQ (Sheng, 2008). The impact of such a violation could mean that the risk of a Type 1 or 2 error was increased (Sheng, 2008), meaning that the statistical inferences made about the impact of self-management and interdependent group contingency on generalised behaviour in this research, may be inaccurate (Shadish et al., 2002). A way to counter this violation would be to randomly allocate individuals to conditions (Shadish et al., 2002) and this would be recommended for future research looking at the impact of these approaches on generalised behaviour. Alternatively, generalised behaviour could have been measured using whole-class observation with visual inspection used to analyse this data, without use of the SDQ.

Three observations were conducted at each time point for each class, as it was felt that conducting only one observation per time point, would give an unreliable estimate of off-task and disruptive behaviour. Given that there were multiple observations conducted, visual inspection was chosen to analyse this data, as more could be analysed from the data than just the mean or 'level', such as overlapping data points, to reach a conclusion around whether there had been a change in behaviour. However, a full visual inspection as described by Kratochwill et al. (2010) (see Section 4.2.2) could not be conducted. For

instance, trend could not be reliably measured in all cases, with three data points. As such only a partial visual inspection could be conducted, giving limited analyses.

5.4.5. **The approaches**

Fidelity checks highlighted that the pupils in Classes 1 and 2 understood the approaches and were able to follow them. Given that there are a number of ways in which to implement self-management and interdependent group contingency (Chafouleas et al., 2012; Davies & Witte, 2000; Denune et al., 2015; Hoff & Ervin, 2013), this research can only illustrate how effective this particular type of self-management (which used a 0-4 rating scale), and this type of interdependent group contingency (which used a 50 point goal with a 0-4 point scale) was *in this particular context*. This research cannot answer how effective other types of self-management or interdependent group contingency approaches are. Further research would have to be conducted in other settings to gain a general picture of the effectiveness of both approaches when conducted in different ways.

5.4.6. **Implementation and fidelity**

A significant limitation in this research is the degree to which the approaches were implemented. It was recommended that the approach be implemented daily and this was agreed with the teachers in advance. However, due to staff absence and activities taking place in school that prevented the approaches from being conducted every day, during Phase A, self-management was only conducted four times. As a result, the findings, which suggest that self-management may not have led to generalised changes in behaviour or a reduction in disruptive behaviour, may in fact have been due to it not having been conducted as often as necessary. Furthermore, the reduction in off-task behaviour may have been greater had the approach been conducted daily. On reflection, it would have been beneficial to have conducted a pilot, which would have highlighted that it was not possible to run this approach every afternoon. Had this been highlighted, Phase A could

have been extended so that the class had enough practise to better internalise the skills of self-management.

It was suggested prior to the research that the teachers in Classes 1, 2 and 3 keep a log of when they implement the approach (or in the case of Class 3, when the class *were* reminded of the rules), as well as anything that may have prevented these approaches from being implemented on a particular day, however none of the teachers found it possible to do this in a formal way. In retrospect, it may have been easier to supply the teachers with a calendar that they could pin onto a wall, and simply tick the days when the approach was implemented.

Although the researcher did implement a number of strategies that the literature has identified support implementation, such as creating buy-in through involving executives, building relationships, producing appropriate training materials and providing on-going support (Powell et al., 2012), a number of other relevant factors were not considered and these may account for why the approaches were not implemented as often as expected. For instance, management were not trained on the approaches in order that they could support the teachers with implementation (Forman & Barakat, 2011), through regular coaching, assessment and feedback (Fixsen et al., 2005).

Damschroder et al., (2009) outlined five factors that affect the successful implementation of an intervention (see Table 2.1). These factors are intervention characteristics, outer settings, inner settings, individual characteristics and the process of implementation. Table 5.1 outlines the factors within this research which may have contributed to difficulties with implementing self-management and interdependent group contingency regularly. In view of these factors, actions that future researchers and trainers could take when planning to implement these approaches are also outlined in Table 5.1.

Factors	Examples	Remediation
Intervention characteristics	• The staff in school may have perceived that the approaches would be ineffective / did not believe that they could improve behaviour.	• More time should be spent at the recruitment or training stage, to outline the evidence-base for these approaches with in-depth information given on which behaviours specifically have improved in the research, as a result of using these approaches.
	• The teachers implementing the approaches initially raised concerns that it was difficult to remember the stages of self-management and to stop particular activities the moment the buzzer went off.	• Pilot the approaches in other classrooms to see what the pragmatic difficulties may be and make adjustments before starting future research using these approaches. • Simplify the self-management procedure, for instance by using a thumbs up or thumbs down approach (Briesch et al., 2013). • Investigate whether all of the stages of self-management used in this study were necessary (see Section 5.6) and eliminate any that seem less important.
Outer settings	• The teachers may have experienced pressures and difficulties outside of school, such as illness or other personal difficulties which may have affected their readiness and ability to engage with these approaches.	• Ensure that the approaches are simple enough for the teacher to implement without very many pragmatic difficulties. • Ensure that the approaches are simple enough to be implemented by cover teachers or the teaching assistant within the classroom, if there is teacher absence. • Ensure other staff members who may cover the class when the teacher is absent, receive the training.

Factors	Examples	Remediation
Inner settings	• The teachers may have felt that there was little advantage to them participating in this project. They may have struggled to perceive meaningful improvement, or as much improvement as they had hoped for.	• Engage with senior management to develop in-house incentives and rewards for the teachers to participate in this project, beyond the hoped for improvements with behaviour in the classroom.
	• The teachers may have felt under pressure with large workloads and other required school tasks, which affected how prepared they felt to run the approach each day. This is particularly with self-management which required rating sheets to be printed and trimmed ahead of time.	• Work with senior management to ensure that these approaches are a priority for the teachers and that as such, they are afforded the time or support to prepare the necessary materials. Where possible, senior management should be encouraged to reduce other pressures on the teachers, to provide capacity to take on something new. Time should be spent discussing with management, the advantages of allowing these approaches to be implemented regularly and to fidelity.
Individual characteristics	• Though the teachers consented to take part, they may in reality have not felt enthusiastic about the extra work required and may have felt reluctant to implement it for reasons such as the complexity of self-management.	As described above in more detail: • Simplify the approaches • Create buy-in by highlighting the evidence-base and documented behavioural improvements following the proper implementation of these approaches • Work with senior management to reduce other pressures on the teachers and provide incentives and rewards

Factors	Examples	Remediation
Process of implementation	• Enough time was not spent in understanding the pressure on the teachers and the potential obstacles they faced in implementing these approaches.	• More time needs to be given to doing a needs analysis within the school and understanding what pressures the teachers face. Time also needs to be spent observing and problem-solving particular obstacles that the teachers could face in implementing these approaches and action should be taken to ensure the approaches are as manageable as possible. An example is through conducting action research, by which changes are made until the approaches are feasible to implement and effective within the school context (see Section 5.6).

Table 5.1: A table outlining the factors that may have affected implementation of the approaches in this research, as well as ideas for remediation, based on the five areas that affect successful implementation, as outlined by Damschroder et al., (2009).

Other barriers may have included teacher perceptions of the relative advantage of these approaches over other techniques, the cost of implementing the approaches (i.e. having to stop the class at regular intervals with self-management which disrupts the lesson, or the time it takes to photocopy and trim new rating sheets), external pressures of working in a busy school, the perceived priority of these approaches compared to other priorities, and the implementers' readiness for change (Damschroder et al., 2009). Implementation science literature suggests that planning strategies to overcome such potential barriers from the start may have increased the likelihood of better implementation (Damschroder et al., 2009). Such strategies include conducting a thorough assessment of what the potential barriers may be (i.e. planning, preparation and assessment time, library slot, assembly, PE), providing incentives to the teachers and implementing the approaches school-wide to enable teachers to meet with and shadow other implementers (Powell et al., 2012). These are strategies the researcher could have considered beforehand. Briesch (2013) implemented a simpler version of self-management in which the pupils

rated their behaviour using their thumbs. Streamlining the approach in this way may have supported its implementation (Briesch, Briesch, et al., 2015).

In order to reduce the risk that a lack of fidelity to the intervention impacted the findings, fidelity checks were conducted by the researcher and four PAs, following training. On the whole, the approaches were conducted to fidelity and in the control classes, there was no evidence of diffusion of treatment (see appendix 8.16). It is possible, however, that diffusion of treatment did occur but was not observed on the days when the teachers of the control classes knew they were being observed. A fidelity check was always conducted at the start of a phase, however, due to unforeseen circumstances within schools it was not always possible to conduct them half way through, with appointments often cancelled or rearranged. As such, it is possible that adherence to the exact procedures may have reduced as the weeks went on, within a phase, making the approaches less effective and impacting the final results. On reflection, the researcher could have asked that lessons be video recorded so that fidelity checks could still be conducted. Also, as previously suggested, targeting a morning lesson would have increased the opportunities to do a fidelity check due to a greater likelihood that the approaches would have been running daily.

5.4.7. **Inter-observer agreement**

Before conducting IOA checks, the PAs received training that included discrepancy discussions (Yoder & Symons, 2010), however, there was insufficient time to do much practice before collecting the data. This may have impacted on the reliability of the PAs' scoring, although these scores did not vary greatly from the researcher's own scores. Having a number of PAs conducting IOA checks may have added an extra 'person' variable to the outcomes of the checks, although it may also have prevented *observer drift,* with the increased likelihood that each PA would adhere closely to the on-task, off-task and disruptive criteria. The PAs knew which class (experimental or control) they were observing. On reflection, it would have been beneficial to have them be

'blind' observers who would not have been swayed by knowing which group they were watching, however, this was not possible as often the PAs were required to support with fidelity checks, at the risk of some not being carried out at all. The PAs' own work schedules and the changing class timetables meant that it was not possible to assign two to conduct only fidelity checks and two to conduct only IOA checks.

5.4.8. Extraneous variables associated with real world research in schools

A significant challenge faced in this research was trying to maintain control over extraneous variables. The research was designed such that all classes' data would be collected during the same two weeks at each time point. However, due to unforeseen circumstances in which the teacher in Class 2 was absent for almost three weeks in the run up to the Christmas holidays, the Time 2 and Time 3 data for this class was collected seven weeks after it had been collected from the other three classes. As such, *maturation* could no longer be controlled for, as the improvements observed in Class 2's off-task and disruptive behaviour at Time 2 may have also been observed in the control classes and in Class 1, had their data also have been collected later. Time 2 data in the other three classes was collected in two of the three weeks leading up to the Christmas holidays, when off-curriculum and unstructured activities were taking place. This may have led to increased off-task or disruptive behaviour being observed than if the observations had been conducted at a different time. The improvements observed in Class 2 after Phase A may simply have been due to maturation or because the data was collected mid-term rather than at the end of the term. Furthermore, having a supply teacher may have negatively impacted on the behaviour of the class, which was not something that the other classes experienced, highlighting *history* as a risk to the internal validity of these findings. This may account for the recorded increases in behavioural problems and hyperactivity and concentration problems at Time 2.

Initially, the researcher planned to observe the classes at each time point such that the time of day and order they were observed in was

counterbalanced. This was to control for time-of-day factors. However in practice, this was not possible due to timetable clashes (with PPA, assembly, PE and library time). In the end, the teachers had to inform the researcher of when they would be delivering appropriate lessons. This meant that sometimes, for some classes, the three observations at a time point were collected on consecutive days. For other classes, there may have been a week long gap between two observations. As such, the observations were not collected at regular intervals, potentially affecting the outcomes, depending on what activities were happening on a particular week. It would have been beneficial if the classes could have been observed in the same week or on the same days.

Also, in the real world it was not possible to control for other variables which are likely to have impacted the behaviour observed in class, such as difficulties that the pupils had experienced in the playground that afternoon. Additionally, the pupils may have felt more tired towards the end of the week. Before and after Christmas, many pupils in the school and in the observed classes experienced a sickness bug; some were sent home. This could have impacted levels of engagement in lessons.

Finally, due to long term teacher absence, Time 3 data could not be collected from Class 2 (IGC+SM), which limited the data available to answer research question 4, and to observe whether difficulties with hyperactivity and concentration were maintained, reduced or increased. Had another school been recruited as well, with more classes implementing these approaches, the impact of mortality could have been reduced.

Given the particular limitations of this research, the findings and conclusions should be interpreted with caution. These limitations are likely to impact the strength of the causal link between the classroom management approaches and the behaviours measured. Despite best efforts to control for extraneous variables, compromises had to be made, which introduced threats to the reliability and internal validity of the data.

5.5. Implications of the findings

This research aimed to evaluate the efficacy of two approaches to classroom management, for improving engagement and behaviour in classes. The implications of these findings for schools and EPs are discussed here.

5.5.1. Implications for schools

The findings which examined the impact of the approaches in target lessons found that interdependent group contingency reduced both off-task and disruptive behaviour, whereas self-management only reduced off-task behaviour (although to a greater degree than interdependent group contingency). The implications of this for schools are that where classes experience mainly off-task (but not disruptive) behaviours, self-management could be most beneficial. If however classes display off-task *and* disruptive behaviours, interdependent group contingency may be a more suitable choice. Furthermore, interdependent group contingency may be more appropriate for an active lesson such as P.E., where it would be difficult to complete self-rating sheets.

The literature indicated that combining these two approaches could be beneficial (Freeman & Dexter-Mazza, 2004; Mitchem et al., 2001; Mooney et al., 2005), however the findings of this research did not support this; combining the approaches had no further impact on engagement behaviours than self-management alone. As such, schools should be cautious about using the combined approach; it may be more efficient and easier to use only one. Teachers should select carefully which approach is most appropriate based on the need in their class.

The findings also suggested that neither approach positively impacted behaviour more generally, outside of the target lesson. The implication of this is that in order to facilitate improvements in general behaviour, the approaches may need to be implemented in *all* lessons, which could be difficult to sustain. It may instead be beneficial to only target lessons that cause most concern.

The simplicity of the approaches suggests that unqualified teachers or teaching assistants could also be trained on these approaches to ensure they are administered even when the class teacher is absent. Also, these approaches could be beneficial to use in settings and classes where there may be greater off-task and disruptive behaviour, such as in Pupil Referral Units. These approaches have been found to be effective on such populations in the US (Denune et al., 2015).

5.5.2. Implications for Educational Psychologists

The findings from this study suggest that these approaches could be effective with a UK population. Added to the US evidence base, EPs could recommend these approaches to their schools, however consideration must be given to the most appropriate way to train staff. Self-management and interdependent group contingency were not implemented as regularly and robustly as indicated in the literature, however these challenges are not uncommon in real world research. EPs need to have an awareness of the implementation science literature when planning and delivering training of these and other approaches. The literature outlines strategies that EPs could consider, to support implementation (see Section 2.4.4), such as extending the training to management as well as to teaching staff, so that schools can develop in-house support networks for these approaches at a strategic level (Fixsen et al., 2005; Forman & Barakat, 2011; Powell et al., 2012). Furthermore, EPs have a role in providing coaching for the successful integration of these procedures in the setting.

This research has highlighted key theories and evidence around the topic of engagement, which would suggest that schools should promote pupil autonomy, competence (Urdan & Schoenfelder, 2006), and feelings of relatedness (Deci & Ryan, 1985; Reeve, 2012; Ryan & Deci, 2000), by promoting positive interactions between teachers and pupils (Garon-Carrier et al., 2015; Jang et al., 2010; Lavigne et al., 2007; F. Mitchell et al., 2015; Shih, 2008; Sparks et al., 2015; Standage et al., 2005; Stroet et al., 2013; Van den Berghe et al., 2016), and improving the quality of teaching (Fredricks et al., 2004), rather than focusing on punitive approaches which are often adopted in schools' behaviour

policy. EPs have a role in advocating for these underlying principles and theories to be adopted by head teachers, with a view to impacting schools at a strategic policy and practice level.

5.6. **Future research**

Discussion of the findings from this research has led to the identification of a number of avenues for further research. Given the shortcomings identified within this study, it would be useful for future research to replicate the current design or to implement a multiple baseline across classes design. A design such as this would allow for a full visual inspection as well as more reliable trends to be identified. Furthermore, as each class could be given the approaches at different times, if all showed improvements only when the intervention(s) were implemented, this would strengthen the causal link. In addition, this design would allow more observations to be carried out, leading to more reliable data for each phase. Replication of the current study's research questions (whether using the current design or a multiple baseline design) would provide further evidence toward either accepting or rejecting the null hypothesis, while increasing external validity. Beyond this, future studies might explore other, related questions such as:

- ***Does implementing interdependent group contingency lead to greater general difficulties with behaviour and hyperactivity and concentration?***

This study raised questions regarding whether interdependent group contingency may lead to greater general behaviour difficulties in school. A replication of the current study would help to explore this relationship further, to identify whether these findings were erroneous or causally linked.

- ***What is the impact of varying the target goal for interdependent group contingency?***

It may be that the limited benefits observed from implementing interdependent group contingency were due to the goal being too

ambitious and not achievable in a short time frame, thereby reducing motivation due to appearing unachievable. Future research could compare the differential impact of implementing interdependent group contingency with 30, 40 or 50 points as the target.

- ***Which approach to self-management is the most effective?***

As discussed in Chapter 2, there are a number of ways in which self-management could be implemented; there are no set procedures. The limited impact of self-management in this research may not be generalisable to other forms of self-management. As such, it would be useful for future research to compare the impact of different approaches. For instance, one could explore whether it is more effective to have pupils rate whether or not they are following a rule *at the moment the buzzer goes off* rather than by reflecting on the preceding 15 minutes.

- ***What impact does self-management have on off-task and disruptive behaviour if it is implemented daily, compared to weekly or twice weekly?***

The findings of this research indicated that in Phase A of the study, pupils who received self-management reduced their off-task behaviours over and above all other groups. This was despite the fact that the approach had only been administered eight times. As discussed previously, this may have been as a result of extraneous variables, however, it may indicate that self-management does not need to be implemented daily for there to be a positive impact in the target lesson. As such, it would be of great interest to explore the impact of administering self-management weekly, twice weekly and daily, to see whether implementing it more often leads to greater benefits. If the findings suggest that there is no additional benefit to implementing the approach daily compared to twice weekly, this may make it more manageable for teachers to implement.

- ***What are the supports and barriers to implementing self-management and/or interdependent group contingency?***

Implementation of the approaches in this research was low; the approaches were not applied as often as one might ideally expect. Some of the barriers to its implementation were identified in conversation with the teachers, however, explorative research to identify additional barriers and supports to implementation would be valuable, given that much research indicates these approaches improve behaviour (Chafouleas et al., 2012; Davies & Witte, 2000; Denune et al., 2015; Glynn et al., 1973; Hoff & Ervin, 2013). Furthermore, action research to identify the conditions under which these approaches can be implemented successfully and frequently, would be valuable.

- ***What impact does self-management and interdependent group contingency have on the pupils who experience the greatest difficulty in following classroom expectations?***

This research focused on the impact of these approaches on the class as a whole and the impact was small at best. This may be because no improvement could be observed in the large number of pupils in every class, who were mostly engaged and following the rules from the beginning. The approaches may have had a substantial impact on those pupils who found it difficult to follow the rules at Time 1 but this was not explored. As such, future research could explore the impact on just those target pupils. Alternatively, this study could be replicated in more challenging classrooms, where more disruption is observed; for instance, within a Pupil Referral Unit.

- ***What are the essential aspects of self-management?***

The fidelity checks conducted for this research highlighted that when self-management was implemented in Class 1, the teacher at times did not remind the class of the rules at the beginning and at times forgot to ask the pupils to share their ratings with each other. Nevertheless, the findings indicated that off-task behaviour improved during Phase A. As such, it may be that reminding pupils of the rules and/or sharing their ratings with a peer are unnecessary steps which do not lead to additional benefits; it may be that simply stopping, reflecting and self-rating is sufficient. Future research should explore this by comparing a number of variations of the self-management procedure implemented

in this research. If it appears that daily rule reminders and peer feedback is unnecessary, the simplified procedure may improve fidelity and the frequency of its implementation in the classroom.

- ***Which is most successful: implementing self-management and interdependent group contingency or supporting teachers to be more autonomy- and competence-supportive in lessons?***

The literature discussed in Section 2.3.3 highlighted other relevant research areas for engagement and disruptive behaviour, which focused on teacher-pupil interactions (Evertson & Weinstein, 2006; Hajdukova et al., 2014; Marsh, 2012), where pupil autonomy and competence with a task were promoted (Garon-Carrier et al., 2015; Jang et al., 2010; Lavigne et al., 2007; F. Mitchell et al., 2015; Shih, 2008; Sparks et al., 2015; Standage et al., 2005; Stroet et al., 2013; Van den Berghe et al., 2016), among a range of other factors. It would be valuable to explore which approach leads to better engagement in lessons with less disruption; focusing on self-regulation and behaviour or intervening with classroom interactions and the tasks presented within a lesson.

5.7. Researcher reflections

The research questions of this study aimed to evaluate the effectiveness of two whole-class approaches to classroom management with a UK population. This was a useful undertaking as the current evidence in support of these approaches has come from international populations, however disengagement and disruption is of concern in UK schools as well. With such a strong US evidence base, it was important to investigate whether these approaches could also be beneficial for UK populations. Reviewing the literature however, highlighted that within lessons, the level of autonomy-, competence- and relatedness-supportive teacher-pupil interactions (Garon-Carrier et al., 2015; Jang et al., 2010; Lavigne et al., 2007; F. Mitchell et al., 2015; Shih, 2008; Sparks et al., 2015; Standage et al., 2005; Stroet et al., 2013; Van den Berghe et al., 2016) can have a significant impact on engagement. As such, it would have been useful to have investigated

the impact of intervening with context-specific factors such as classroom interactions and task characteristics.

6. Conclusion

6.1. Introduction

In this chapter, the key findings from the research are highlighted and the study's unique contribution is identified.

6.2. Unique contribution of this research

A wealth of international research suggests that implementing interdependent group contingency is an effective whole-class classroom management approach for increasing engagement and reducing disruptive behaviours (Christ & Christ, 2006; Ennis et al., 2016; Hansen & Lignugaris-Kraft, 2005; Hartman & Gresham, 2016; Kelshaw-Levering et al., 2000; Ling & Barnett, 2013; Ling et al., 2011; McKissick et al., 2010; Murphy et al., 2007; Theodore et al., 2004; Williamson et al., 2009). Conversely, the evidence for whole-class self-management is limited, although it has a strong international evidence base as a targeted intervention (DuPaul et al., 2011; Freeman & Dexter-Mazza, 2004; Gureasko-Moore et al., 2006; Holifield et al., 2010; Kern et al., 1994; King-Sears, 2008; S.-H. Lee et al., 2007; Miller et al., 1993; Mitchem et al., 2001; Rafferty, 2012; Rooney et al., 1984).

A review of the literature highlighted that despite much research in support of these approaches, none of the research identified in this study was conducted with a UK population. Furthermore, despite claims that interdependent group contingency and self-management lead to maintained and generalised improvements in behaviour (Ennis et al., 2016; Freeman & Dexter-Mazza, 2004; Mitchem et al., 2001; Mooney et al., 2005), research around interdependent group contingency suggested this was not the case (Kelshaw-Levering et al., 2000; Ling & Barnett, 2013; Ling et al., 2011; Murphy et al., 2007), and other studies discussed in the systematic review in Section 2.5 had not measured this. There also appeared to be limited research evaluating the impact of implementing both approaches together compared to implementing only one (Briesch et al., 2013; Chafouleas et al., 2012; Davies & Witte, 2000; Denune et al., 2015; Hoff & Ervin, 2013). As such,

the present study provided a unique contribution to the literature by examining the efficacy of these two approaches on the behaviour of classes in the UK, by seeking to identify within-lesson and generalised changes in behaviour and by measuring whether there was any change in behaviour observed after one approach was added to the other. Furthermore, this research employed a group design and used standardised measures to analyse statistical significance, which is limited in the literature (Bruhn et al., 2015).

6.3. **Key findings**

The key findings from this research were that:

- Implementing self-management leads to reductions in off-task behaviour but no evidence emerged to suggest that self-management reduces disruptive behaviour in the target lessons.
- Implementing interdependent group contingency leads to reductions in off-task and disruptive behaviour in the target lessons, however with comparatively less reductions in off-task behaviour than self-management.
- No evidence emerged to suggest that implementing both approaches together leads to further reductions in off-task and disruptive behaviour in the target lessons.
- No evidence emerged to suggest that implementing self-management leads to generalised improvements in behavioural and hyperactivity and concentration problems of the class.
- No evidence emerged to suggest that implementing interdependent group contingency leads to general improvements in behaviour, however evidence suggests that this approach may increase the general hyperactivity and concentration problems of the class.

Overall, the findings suggest that both approaches are effective in the lessons in which they are implemented, however, they each impact behaviour in different ways. There also appears to be little obvious advantage to implementing both together and no general improvements in behaviour outside of the target lessons. In light of the

significant limitations and the threats to internal validity and reliability of this research discussed in Section 5.4, these conclusions are tentative and must be interpreted with caution.

The current research provides provisional evidence towards the efficacy of these approaches with a UK population. Future research should seek to address the limitations of the current study, and to undertake further evaluation with a variety of age groups in order to contribute towards building a UK evidence-base for the approaches.

7. References

Agarwal, B. (2003). *Programmed statistics (question-answers)* (2nd ed.). New Delhi, India: New Age International.

Altman, K., & Linton, T. (1971). Operant conditioning in the classroom setting: A review of the research. *The Journal of Educational Research, 64*(6), 277-286.

Amato-Zech, N. A., Hoff, K. E., & Doepke, K. J. (2006). Increasing on-task behavior in the classroom: Extension of self-monitoring strategies. *Psychology in the Schools, 43*(2), 211-221.

Anderson, G. J., Walberg, H. J., & Welch, W. W. (1969). Curriculum effects on the social climate of learning: A new representation of discriminant functions. *American Educational Research Journal, 6*(3), 315-328.

Anderson, N. B. (2006). Evidence-based practice in psychology. *American Psychologist, 61*(4), 271-285.

Appleton, J. J., Christenson, S. L., & Furlong, M. J. (2008). Student engagement with school: Critical conceptual and methodological issues of the construct. *Psychology in the Schools, 45*(5), 369-386.

Ardoin, S. P., & Martens, B. K. (2004). Training children to make accurate self-evaluations: Effects on behavior and the quality of self-ratings. *Journal of Behavioral Education, 13*(1), 1-23.

Association of Teachers and Lecturers. (2010). Over a quarter of education staff have dealt with physical violence by pupils [Press release]. Retrieved from https://www.atl.org.uk/Images/29%20March%202010%20%20ATL%20Behaviour%20Survey%20%20-%20Annual%20Conference%20-%20final.pdf

Axup, T., & Gersch, I. (2008). The impact of challenging student behaviour upon teachers' lives in a secondary school: teachers' perceptions. *British Journal of Special Education, 35*(3), 144-151.

Ayers, H., Clarke, D., & Murray, A. (2000). *Perspectives on behaviour: A practical guide to effective interventions for teachers* (2nd ed.). London, UK: David Fulton Publishers.

Bagley, W. C. (1907). *Classroom management: Its principles and technique*. New York: Macmillan.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hill.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287.

Barlow, D. H., & Hersen, M. (1984). *Single case experimental designs: Strategies for measuring behavior change* (2nd ed.). Oxford, UK: Pergamon Press.

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Boston, MA: Allyn & Bacon

Barrett, P., Davies, F., Zhang, Y., & Barrett, L. (2015). The impact of classroom design on pupils' learning: Final results of a holistic, multi-level analysis. *Building and Environment, 89*, 118-133.

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good behavior game: effects of individual contingencies for group consequences on disruptive behavior in a classroom. *Journal of Applied Behavior Analysis, 2*(2), 119-124.

Barry, L. M., & Haraway, D. L. (2005). Self-management and ADHD: A literature review. *The Behavior Analyst Today, 6*(1), 48-64.

Bednar, R. L., Zelhart, P. F., Greathouse, L., & Weinberg, S. (1970). Operant conditioning principles in the treatment of learning and behavior problems with delinquent boys. *Journal of Counseling Psychology, 17*(6), 492-497.

Blackman, D. E. (1984). The current status of behaviourism and learning theory in psychology. In D. Fontana (Ed.), *Behaviourism and learning theory in education* (pp. 3-14). Edinburgh, UK: Scottish Academic Press Limited.

Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher–pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction, 21*(6), 715-730.

Boekaerts, M., Pintrich, P., & Zeidner, M. (2000). Self-regulation: An introductory overview. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 1-9). San Diego, CA: Academic Press.

Briesch, A. M., Briesch, J. M., & Chafouleas, S. M. (2015). Investigating the usability of classroom management strategies among elementary schoolteachers. *Journal of Positive Behavior Interventions, 17*(1), 5-14.

Briesch, A. M., Hemphill, E., & Daniels, B. (2013). Check your SLANT: Adapting self-management for use as a class-wide intervention. *School Psychology Forum, 7*(2), 29-39.

Briesch, A. M., Hemphill, E. M., Volpe, R. J., & Daniels, B. (2015). An evaluation of observational methods for measuring response to classwide intervention. *School Psychology Quarterly, 30*(1), 37-49.

British Psychological Society. (2009). *Code of ethics and conduct: Guidance published by the ethics committee of the British Psychological Society*. Leicester, UK: British Psychological Society.

British Psychological Society. (2014). *Code of human research ethics*. Leicester, UK: British Psychological Society.

Brophy, J. (2006). History of research on classroom management. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 17-43). Mahwah, MJ: Lawrence Erlbaum Associates.

Brown, A. F. (1966). A perceptual taxonomy of the effective-rated teacher. *The Journal of Experimental Education, 35*(1), 1-10.

Bruhn, A., McDaniel, S., & Kreigh, C. (2015). Self-Monitoring interventions for students with behavior problems: A systematic review of current research. *Behavioral Disorders, 40*(2), 102-121.

Buhs, E. S., Ladd, G. W., & Herald, S. L. (2006). Peer exclusion and victimization: Processes that mediate the relation between peer group rejection and children's classroom engagement and achievement? *Journal of Educational Psychology, 98*(1), 1-13.

Burchard, J., & Tyler, V. (1964). The modification of delinquent behaviour through operant conditioning. *Behaviour Research and Therapy, 2*(2), 245-250.

Burton, K. D., Lydon, J. E., D'alessandro, D. U., & Koestner, R. (2006). The differential effects of intrinsic and identified motivation on well-being and performance: Prospective, experimental, and implicit approaches to self-determination theory. *Journal of Personality and Social Psychology, 91*(4), 750-762.

Caldarella, P., Williams, L., Jolstead, K. A., & Wills, H. P. (2016). Managing Student Behavior in an Elementary School Music Classroom: A Study of Class-Wide Function-Related Intervention Teams. *Update: Applications of Research in Music Education*, 1-8.

Calvert, S. C., & Johnston, C. (1990). Acceptability of treatments for child behavior problems: Issues and implications for future research. *Journal of Clinical Child Psychology, 19*(1), 61-74.

Canter, L. (2010). *Assertive Discipline: Positive behavior management for today's classroom* (4th ed.). Bloomington, IN: Solution Tree Press.

Canter, L., & Canter, M. (1976). *Assertive discipline: A take charge approach for today's educator*. Santa Monica, CA: Lee Canter and Associates.

Carter, K., & Doyle, W. (2006). Classroom management in early childhood and elementary classrooms. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 373-406). Mahwah, NJ: Lawrence Erlbaum Associates.

Cavalier, A. R., Ferretti, R. P., & Hodges, A. E. (1997). Self-management within a classroom token economy for students with learning disabilities. *Research in Developmental Disabilities, 18*(3), 167-178.

Chafouleas, S. M., Sanetti, L. M. H., Jaffery, R., & Fallon, L. M. (2012). An evaluation of a classwide intervention package involving self-management and a group contingency on classroom behavior of middle school students. *Journal of Behavioral Education, 21*(1), 34-57.

Cheon, S. H., & Reeve, J. (2015). A classroom-based intervention to help teachers decrease students' amotivation. *Contemporary Educational Psychology, 40*(1), 99-111.

Christ, T. J., & Christ, J. A. (2006). Application of an interdependent group contingency mediated by an automated feedback device: An intervention across three high school classrooms. *School Psychology Review, 35*(1), 78-90.

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement*. New York, NY: Springer.

Clark, E., Olympia, D. E., Jensen, J., Heathfield, L. T., & Jenson, W. R. (2004). Striving for autonomy in a contingency-governed world: Another challenge for individuals with developmental disabilities. *Psychology in the Schools, 41*(1), 143-153.

Clemens, N. H., Shapiro, E. S., & Seibert, A. L. (2013). Conducting systematic direct classroom observations to define and assess school-related problems. In R. Brown-Chidsey & K. J. Andren (Eds.), *Assessment for intervention: A problem-solving approach* (2nd ed., pp. 157-179). New York, NY: The Guilford Press.

Clunies-Ross, P., Little, E., & Kienhuis, M. (2008). Self-reported and actual use of proactive and reactive classroom management strategies and their relationship with teacher stress and student behaviour. *Educational Psychology, 28*(6), 693-710.

Coe, R. J. (2012). The nature of educational research - exploring the different understandings of educational research. In J. Arthur, M. Waring, R. Coe, & L. V. Hedges (Eds.), *Research methods and methodologies in education* (pp. 5-14). London: Sage Publications Ltd.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London, UK: RoutledgeFalmer.

Concato, J., Shah, N., & Horwitz, R. I. (2000). Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine, 342*(25), 1887-1892.

Coogan, B. A., Kehle, T. J., Bray, M. A., & Chafouleas, S. M. (2007). Group contingencies, randomization of reinforcers, and criteria for reinforcement, self-monitoring, and peer feedback on reducing inappropriate classroom behavior. *School Psychology Quarterly, 22*(4), 540-556.

Corpus, J. H., & Wormington, S. V. (2014). Profiles of intrinsic and extrinsic motivations in elementary school: A longitudinal analysis. *The Journal of Experimental Education, 82*(4), 480-501.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage Publications Inc.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage Publications.

Cross, M. (2004). *Children with emotional and behavioural difficulties and communication problems: There is always a reason*. London, UK: Jessica Kingsley Publishers.

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. London, UK: Sage Publications.

Dalton, T., Martella, R. C., & Marchand-Martella, N. E. (1999). The effects of a self-management program in reducing off-task behavior. *Journal of Behavioral Education, 9*(3-4), 157-176.

Damschroder, L. J., Aron, D. C., Keith, R. E., Kirsh, S. R., Alexander, J. A., & Lowery, J. C. (2009). Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. *Implementation science, 4*(50), 1-15.

Dancey, C. P., & Reidy, J. (2007). *Statistics without maths for psychology* (4th ed.). Essex, UK: Pearson Education Ltd.

Dart, E. H., Radley, K. C., Battaglia, A. A., Dadakhodjaeva, K., Bates, K. E., & Wright, S. J. (2016). The classroom password: A class-wide intervention to increase academic engagement. *Psychology in the Schools, 53*(4), 416-431.

Davies, S., & Witte, R. (2000). Self-management and peer-monitoring within a group contingency to decrease uncontrolled verbalizations of children with attention-deficit/hyperactivity disorder. *Psychology in the Schools, 37*(2), 135-147.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627-668.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Denune, H., Hawkins, R., Donovan, L., Mccoy, D., Hall, L., & Moeder, A. (2015). Combining self-monitoring and an interdependent group contingency to improve the behavior of sixth graders with EBD. *Psychology in the Schools, 52*(6), 562-577.

Department for Education. (2012). *Pupil behaviour in schools in England* (DFE-RR218). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/184078/DFE-RR218.pdf:

Department for Education. (2014). *Permanent and fixed-period exclusions in England: 2009 to 2014*. Retrieved from https://www.gov.uk/government/collections/statistics-exclusions.

Department for Education. (2016). *Behaviour and discipline in schools: Advice for headteachers and school staff*. London, UK: Crown Copyright.

Department for Education. (2017). *Creating a culture: How school leaders can optimise behaviour*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/602487/Tom_Bennett_Independent_Review_of_Behaviour_in_Schools.pdf.

Department for Education and Employment. (1998). Excellent in research on schools. In. London, UK: The Stationery Office.

Department for Education and Science. (1989). *Discipline in schools [The Elton Report]*. London: HMSO.

Department for Education and Skills. (2005). *Learning behaviour: The report of the practitioner group on school behaviour and discipline [The Steer Report]Learning behaviour: The report of the practitioner group on school behaviour and discipline [The Steer Report]*. London: DfES.

Department for Education and Skills. (2012). *Practical approaches to behaviour management in the classroom: A handbook for classroom teachers in primary schools*. Cardiff, UK: Welsh Government.

Department of Health. (1998). A first class service: Quality in the new NHS. In. London, UK: The Stationery Office.

Deutsch, M. (1949). A theory of co-operation and competition. *Human Relations, 2*(2), 129-152.

Dewar, D., Plotnikoff, R., Morgan, P., Okely, A., Costigan, S., & Lubans, D. (2013). Testing social-cognitive theory to explain physical activity change in adolescent girls from low-income communities. *Research Quarterly For Exercise and Sport, 84*(4), 483-491.

Dickson, A., Cherry, M. G., & Boland, A. (2014). Carrying out a systematic review as a master's thesis. In A. Boland, M. G. Cherry, & A. Dickson (Eds.), *Doing a systematic review: A student's guide* (pp. 1-16). London: Sage Publications.

Dixon, J. C., Singleton Jr, R. A., & Straits, B. C. (2016). *The process of social research*. New York, NY: Oxford University Press.

Dundar, Y., & Fleeman, N. (2014). Developing my search strategy and applying inclusion criteria. In A. Boland, M. G. Cherry, & R. Dickson (Eds.), *Doing a systematic review: A student's guide* (pp. 35-59). London: Sage Publications.

DuPaul, G. J., Weyandt, L. L., & Janusis, G. M. (2011). ADHD in the classroom: Effective intervention strategies. *Theory into Practice, 50*(1), 35-42.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82*(1), 405-432.

Ennis, C. R., Blair, K.-S. C., & George, H. P. (2016). An evaluation of group contingency interventions: The role of teacher preference. *Journal of Positive Behavior Interventions, 18*(1), 17-28.

Evertson, C. M., & Weinstein, C. S. (2006). Classroom management as a field of inquiry. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 3-15). Mahwah, NJ: Lawrence Erlbaum Associates.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London, UK: Sage Publications.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London, UK: Sage Publications Ltd.

Finn, J. D., & Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 97-131). New York, NY: Springer.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Tampa, FL: National Implementation Research Network.

Flanders, N. A. (1961). Analyzing Teacher-Behavior. *Educational Leadership, 19*(3), 173-200.

Flower, A., McKenna, J. W., Bunuan, R. L., Muething, C. S., & Vega, R. (2014). Effects of the Good Behavior Game on challenging behaviors in school settings. *Review of Educational Research, 84*(4), 546-571.

Forman, S. G., & Barakat, N. M. (2011). Cognitive-behavioral therapy in the schools: Bringing research to practice through effective implementation. *Psychology in the Schools, 48*(3), 283-296.

Forman, S. G., Shapiro, E. S., Codding, R. S., Gonzales, J. E., Reddy, L. A., Rosenfield, S. A., . . . Stoiber, K. C. (2013). Implementation science and school psychology. *School Psychology Quarterly, 28*(2), 77-100.

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

Frederickson, N., & Cline, T. (2015). *Special educational needs, inclusion and diversity* (3rd ed.). Berkshire, UK: McGraw-Hill Education.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Freeman, K. A., & Dexter-Mazza, E. T. (2004). Using self-monitoring with an adolescent with disruptive classroom behavior preliminary analysis of the role of adult feedback. *Behavior Modification, 28*(3), 402-419.

Froiland, J. M., & Oros, E. (2014). Intrinsic motivation, perceived competence and classroom engagement as longitudinal

predictors of adolescent reading achievement. *Educational Psychology, 34*(2), 119-132.

Garon-Carrier, G., Boivin, M., Guay, F., Kovas, Y., Dionne, G., Lemelin, J.-P., . . . Tremblay, R. E. (2015). Intrinsic motivation and achievement in mathematics in elementary school: A longitudinal investigation of their association. *Child Development, 87*(1), 165-175.

Gleitman, H., Fridlund, A. J., & Reisberg, D. (2004). *Psychology* (6th ed.). New York: W. W. Norton & Company.

Glynn, E. L., Thomas, J. D., & Shee, S. M. (1973). Behavioral self-control of on-task behavior in an elementary classroom. *Journal of Applied Behavior Analysis, 6*(1), 105-113.

Gomm, R. (2008). *Social research methodology: A critical introduction* (2nd ed.). Hampshire, UK: Palgrave Macmillan.

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 & Adolescent Psychiatry, 40*(11), 1337-1345.

Goodman, R., & 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.

Gordon, A. (2001). School exclusions in England: Children's voices and adult solutions? *Educational Studies, 27*(1), 69-85.

Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), 213-228.

Graham-Day, K. J., Gardner III, R., & Hsin, Y.-W. (2010). Increasing on-task behaviors of high school students with attention deficit hyperactivity disorder: Is it enough? *Education and Treatment of Children, 33*(2), 205-221.

Gray, D. E. (2014). *Doing research in the real world* (3rd ed.). London, UK: Sage Publications Ltd.

Greene, J., & D'Oliveira, M. (2005). *Learning to use statistical tests in psychology* (3rd ed.). Berkshire, UK: Open University Press.

Greenhalgh, J., & Brown, T. (2014). Quality assessment: Where do I begin? In A. Boland, M. G. Cherry, & R. Dickson (Eds.), *Doing a systematic review: A student's guide* (pp. 61-83). London: Sage Publications.

Gu, Q., & Day, C. (2013). Challenges to teacher resilience: Conditions count. *British Educational Research Journal, 39*(1), 22-44.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 105-117). Thousand Oaks, CA: Sage.

Gureasko-Moore, S., Dupaul, G. J., & White, G. P. (2006). The effects of self-management in general education classrooms on the organizational skills of adolescents with ADHD. *Behavior Modification, 30*(2), 159-183.

Hajdukova, E. B., Hornby, G., & Cushman, P. (2014). Pupil–teacher relationships: Perceptions of boys with social, emotional and behavioural difficulties. *Pastoral Care in Education, 32*(2), 145-156.

Hansen, S. D., & Lignugaris-Kraft, B. (2005). Effects of a dependent group contingency on the verbal interactions of middle school students with emotional disturbance. *Behavioral Disorders, 30*(2), 170-184.

Harlacher, J. E. (2015). *Designing effective classroom management*. Bloomington, IN: Marzano Research.

Hart, R. (2010). Classroom behaviour management: Educational psychologists' views on effective practice. *Emotional and Behavioural Difficulties, 15*(4), 353-371.

Hartman, K., & Gresham, F. (2016). Differential effectiveness of interdependent and dependent group contingencies in reducing disruptive classroom behavior. *Journal of Applied School Psychology, 32*(1), 1-23.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Haydn, T. (2012). *Managing pupil behaviour: Improving the classroom atmosphere* (2nd ed.). New York: Routledge.

Hayenga, A. O., & Corpus, J. H. (2010). Profiles of intrinsic and extrinsic motivations: A person-centered approach to motivation and achievement in middle school. *Motivation and Emotion, 34*(4), 371-383.

Health Professions Council. (2008). *Standards of conduct, performance and ethics*. London, UK: Health Professions Council.

Herrera, M., & Little, E. (2005). Behaviour problems across home and kindergarten in an Australian sample. *Australian Journal of Educational & Developmental Psychology, 5*, 77-90.

Hirsch, S. E., Healy, S., Judge, J. P., & Lloyd, J. W. (2016). Effects of an interdependent group contingency on engagement in physical education. *Journal of Applied Behavior Analysis, 49*(4), 975-979.

Hoff, K., & Ervin, R. (2013). Extending self-management strategies: The use of a classwide approach. *Psychology in the Schools, 50*(2), 151-164.

Holifield, C., Goodman, J., Hazelkorn, M., & Heflin, L. J. (2010). Using self-monitoring to increase attending to task and academic accuracy in children with autism. *Focus on Autism and Other Developmental Disabilities, 25*(4), 230-238.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*(2), 65-70.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165-179.

Houser, R. A. (2015). *Counseling and educational research: Evaluation and application* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Hulac, D. M., & Benson, N. (2010). The use of group contingencies for preventing and managing disruptive behaviors. *Intervention in School and Clinic, 45*(4), 257-262.

Imai, K. (2017). *Quantitative social science: An introduction*. Princeton University Press: Princeton, NJ.

Jang, H., Reeve, J., & Deci, E. L. (2010). Engaging students in learning activities: It is not autonomy support or structure but autonomy support and structure. *Journal of Educational Psychology, 102*(3), 588-600.

Johnson, D. W. (2003). Social interdependence: Interrelationships among theory, research, and practice. *American Psychologist, 58*(11), 934-945.

Johnson, D. W., & Johnson, F. P. (1994). *Joining together: Group theory and group skills* (5th ed.). Needham Heights, MA: Allyn and Bacon.

Johnson, D. W., & Johnson, R. T. (2005). New developments in social interdependence theory. *Genetic, Social, and General Psychology Monographs, 131*(4), 285-358.

Johnson, D. W., & Johnson, R. T. (2006). Conflict resolution, peer mediation and peacemaking. In C. M. Evertson & C. S. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 803-832). Mahwah, NJ: Lawrence Erlbaum Associates.

Johnson, D. W., & Johnson, R. T. (2008). Social interdependence theory and co-operative learning: The teacher's role. In R. M. Gillies, A. F. Ashman, & J. Terwel (Eds.), *The teacher's role in implementing co-operative learning in the classroom* (pp. 9-37). New York: Springer.

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin, 89*(1), 47-62.

Johnson, T., Stoner, G., & Green, S. (1996). Demonstrating the experimenting society model with classwide behavior management interventions. *School Psychology Review, 25*(2), 199-214.

Kadzin, A. E. (2003). *Research design in clinical psychology* (4th ed.). Boston, MA: Allyn & Bacon.

Kehle, T. J., Bray, M. A., Theodore, L. A., Jenson, W. R., & Clark, E. (2000). A multi-component intervention designed to reduce disruptive classroom behavior. *Psychology in the Schools, 37*(5), 475-481.

Kelly, B. (2012). Implementation science for psychology in education. In B. Kelly & D. F. Perkins (Eds.), *Handbook of implementation*

science for psychology in education (pp. 3-12). New York, NY: Cambridge University Press.

Kelshaw-Levering, K., Sterling-Turner, H. E., Henry, J. R., & Skinner, C. H. (2000). Randomized interdependent group contingencies: Group reinforcement with a twist. *Psychology in the Schools, 37*(6), 523-533.

Kern, L., Dunlap, G., Childs, K. E., & Clarke, S. (1994). Use of a classwide self-management program to improve the behavior of students with emotional and behavioral disorders. *Education and Treatment of Children, 17*(3), 445-458.

King-Sears, M. E. (2008). Using teacher and researcher data to evaluate the effects of self-management in an inclusive classroom. *Preventing School Failure: Alternative Education for Children and Youth, 52*(4), 25-36.

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health, 74*(7), 262-273.

Koka, A., & Hagger, M. S. (2010). Perceived teaching behaviors and self-determined motivation in physical education: A test of self-determination theory. *Research Quarterly for Exercise and Sport, 81*(1), 74-86.

Kratochwill, T., Hitchcock, J., Horner, R., Levin, J. R., Odom, S., Rindskopf, D., & Shadish, W. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Ladd, G. W., & Dinella, L. M. (2009). Continuity and change in early school engagement: Predictive of children's achievement trajectories from first to eighth grade? *Journal of Educational Psychology, 101*(1), 190-206.

Lane, J. D., & Gast, D. L. (2014). Visual analysis in single case experimental design studies: Brief review and guidelines. *Neuropsychological Rehabilitation, 24*(3-4), 445-463.

Lannie, A. L., & McCurdy, B. L. (2007). Preventing disruptive behavior in the urban classroom: Effects of the good behavior game on

student and teacher behavior. *Education and Treatment of Children, 30*(1), 85-98.

Lavigne, G. L., Vallerand, R. J., & Miquelon, P. (2007). A motivational model of persistence in science education: A self-determination theory approach. *European Journal of Psychology of Education, 22*(3), 351-369.

Lee, S.-H., Simpson, R. L., & Shogren, K. A. (2007). Effects and implications of self-management for students with autism: A meta-analysis. *Focus on Autism and Other Developmental Disabilities, 22*(1), 2-13.

Lee, W., Lee, M.-J., & Bong, M. (2014). Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemporary Educational Psychology, 39*(2), 86-99.

Lemos, M. S., & Veríssimo, L. (2014). The relationships between intrinsic motivation, extrinsic motivation, and achievement, along elementary school. *Procedia-Social and Behavioral Sciences, 112*, 930-938.

Lewin, K. (1935). *A dynamic theory of personality: Selected papers*. New York: McGraw-Hill.

Lewin, K., Lippitt, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created "social climates". *The Journal of Social Psychology, 10*(2), 269-299.

Leyden, G., & Miller, A. (1998). Including all our children in mainstream schools and communities: Realising the dream. *Educational Psychology in Practice, 14*(3), 188-193.

Ling, S., & Barnett, D. (2013). Increasing preschool student engagement during group learning activities using a group contingency. *Topics in Early Childhood Special Education, 33*(3), 186-196.

Ling, S., Hawkins, R., & Weber, D. (2011). Effects of a classwide interdependent group contingency designed to improve the behavior of an at-risk student. *Journal of Behavioral Education, 20*(2), 103-116.

Litow, L., & Pumroy, D. K. (1975). A brief review of classroom group-oriented contingencies. *Journal of Applied Behavior Analysis, 8*(3), 341-347.

Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). *Methods in educational research: From theory to practice* (2nd ed.). San Francisco, CA: John Wiley & Sons.

Lovitt, T. C., & Curtiss, K. A. (1969). Academic response rate as a function of teacher-and self-imposed contingencies. *Journal of Applied Behavior Analysis, 2*(1), 49-53.

Madsen Jr, C. H., Becker, W. C., & Thomas, D. R. (1968). Rules, praise, and ignoring: Elements of elementary classroom control. *Journal of Applied Behavior Analysis, 1*(2), 139-150.

Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of School Psychology, 50*(5), 625-654.

Marsh, H. (2012). Relationships for learning: Using pupil voice to define teacher–pupil relationships that enhance pupil engagement. *Management in Education, 26*(3), 161-163.

Martin, J., McCaughtry, N., Flory, S., Murphy, A., & Wisdom, K. (2011). Using social cognitive theory to predict physical activity and fitness in underserved middle school children. *Research Quarterly For Exercise and Sport, 82*(2), 247-255.

Martin, W. E., & Bridgmon, K. D. (2012). *Quantitative and statistical research methods: From hypothesis to results*. San Francisco, CA: Jossey-Bass.

Marzano, R. J. (2003). *What works in schools: Translating research into action*. Alexandria, VA: Association for Supervision and Curriculum Development.

Massey, N. G., & Wheeler, J. J. (2000). Acquisition and generalization of activity schedules and their effects on task engagement in a young child with autism in an inclusive pre-school classroom. *Education and Training in Mental Retardation and Developmental Disabilities, 35*(3), 326-335.

Masten, A. S., Desjardins, C. D., McCormick, C. M., Sally, I., Kuo, C., & Long, J. D. (2010). The significance of childhood competence and problems for adult success in work: A developmental cascade analysis. *Development and Psychopathology, 22*(3), 679-694.

McAllister, L., Stachowiak, J., Baer, D., & Conderman, L. (1968). The application of operant conditioning techniques in a secondary school classroom. *Journal of Applied Behavior Analysis, 2*(4), 277-285.

McDill, E. L., Meyers Jr, E. D., & Rigsby, L. C. (1967). Institutional effects on the academic behavior of high school students. *Sociology of Education, 40*(3), 181-199.

McKissick, C., Hawkins, R. O., Lentz, F. E., Hailley, J., & McGuire, S. (2010). Randomizing multiple contingency components to decrease disruptive behaviors and increase student engagement in an urban second-grade classroom. *Psychology in the Schools, 47*(9), 944-959.

McLaughlin, T. (1976). Self-control in the classroom. *Review of Educational Research, 46*(4), 631-663.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach*. New York: Plenum Press.

Mertens, D. M. (2005). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (2nd ed.). London, UK: Sage Publications.

Miller, L. J., Strain, P. S., Boyd, K., Jarzynka, J., & McFetridge, M. (1993). The effects of classwide self-assessment on preschool children's engagement in transition, free play and small group instruction. *Early Education and Development, 4*(3), 162-181.

Mitchell, F., Gray, S., & Inchley, J. (2015). 'This choice thing really works…' Changes in experiences and engagement of adolescent girls in physical education classes, during a school-based physical activity programme. *Physical Education and Sport Pedagogy, 20*(6), 593-611.

Mitchell, M. M., & Bradshaw, C. P. (2013). Examining classroom influences on student perceptions of school climate: The role of classroom management and exclusionary discipline strategies. *Journal of School Psychology, 51*(5), 599-610.

Mitchem, K. J., & Young, K. R. (2001). Adapting self-management programs for classwide use: Acceptability, feasibility, and effectiveness. *Remedial and Special Education, 22*(2), 75-88.

Mitchem, K. J., Young, K. R., West, R. P., & Benyo, J. (2001). CWPASM: A classwide peer-assisted self-management program for general education classrooms. *Education and Treatment of Children, 24*(2), 111-140.

Mooney, P., Ryan, J. B., Uhing, B. M., Reid, R., & Epstein, M. H. (2005). A review of self-management interventions targeting academic outcomes for students with emotional and behavioral disorders. *Journal of Behavioral Education, 14*(3), 203-221.

Mufti, E., & Peace, M. (2012). *Teaching and learning and the curriculum: A critical introduction*. London, UK: Continuum International Publishing Group.

Mumby, P. J. (2002). Statistical power of non-parametric tests: A quick guide for designing sampling strategies. *Marine Pollution Bulletin, 44*(1), 85-87.

Murphy, K. A., Theodore, L. A., Aloiso, D., Alric-Edwards, J. M., & Hughes, T. L. (2007). Interdependent group contingency and mystery motivators to reduce preschool disruptive behavior. *Psychology in the Schools, 44*(1), 53-63.

Nelson, L. (1963). Changing concepts of teacher leadership. *Educational Leadership, 21*(3), 177-182.

Nicolson, D., & Ayers, H. (2004). *Adolescent problems: A guide for teachers, parents and counsellors*. London, UK: David Fulton Publishers.

Nie, Y., & Lau, S. (2009). Complementary roles of care and behavioral control in classroom management: The self-determination theory perspective. *Contemporary Educational Psychology, 34*(3), 185-194.

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education, 7*(2), 133-144.

Niesyn, M. E. (2009). Strategies for success: Evidence-based instructional practices for students with emotional and behavioral disorders. *Preventing School Failure: Alternative Education for Children and Youth, 53*(4), 227-234.

Ninci, J., Vannest, K., Willson, V., & Zhang, N. (2015). Interrater agreement between visual analysts of single-case data: A meta-analysis. *Behavior Modification, 39*(4), 510-541.

Nock, M. K., Michel, B. D., & Photos, V. I. (2008). Single-case research designs. In D. McKay (Ed.), *Handbook of research methods in abnormal and clinical psychology* (pp. 337-350). Thousand Oaks, CA: Sage Publications.

Norman, G. R., & Streiner, D. L. (2008). *Biostatistics: The bare essentials* (3rd ed.). Ontario, Canada: BC Decker Inc.

Northup, J., Broussard, C., Jones, K., George, T., Vollmer, T. R., & Herring, M. (1995). The differential effects of teacher and peer attention on the disruptive classroom behavior of three children with a diagnosis of attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis, 28*(2), 227-228.

O'Keefe, P. A., & Linnenbrink-Garcia, L. (2014). The role of interest in optimizing performance and self-regulation. *Journal of Experimental Social Psychology, 53*, 70-78.

O'Leary, K., Becker, W., Evans, M., & Saudargas, R. (1969). A token reinforcement program in a public school: a replication and systematic analysis. *Journal of Applied Behavior Analysis, 2*(1), 3-13.

Ofsted. (2005). *The annual report of Her Majesty's Chief Inspector of Schools 2003/04.* London: The Stationery Office.

Oga-Baldwin, W. Q., Nakata, Y., Parker, P., & Ryan, R. M. (2017). Motivating young language learners: A longitudinal model of self-determined motivation in elementary school foreign language classes. *Contemporary Educational Psychology, 49*(1), 140-150.

Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). Teacher classroom management practices: Effects on disruptive or aggressive. *Campbell Systematic Reviews.*

Olswang, L. B., & Prelock, P. A. (2015). Bridging the gap between research and practice: Implementation science. *Journal of Speech, Language, and Hearing Research, 58*(6), S1818-S1826.

Osborne, J. (1968). Free-time as a reinforcer in the management of classroom behavior. *Journal of Applied Behavior Analysis, 2*(2), 113-118.

O'Connor, E. E., Dearing, E., & Collins, B. A. (2011). Teacher-child relationship and behavior problem trajectories in elementary school. *American Educational Research Journal, 48*(1), 120-162.

Palincsar, A. S., & Parecki, A. D. (1995). Important issues related to single-subject experimental research. In S. B. Neuman & S. McCormick (Eds.), *Single-subject experimental research: Applications for literacy* (pp. 137-150). Newark, DE: International Reading Association, Inc.

Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS version 15* (3rd ed.). Maidenhead, UK: Open University Press.

Partington, G. (2001). Student suspensions: The influence on students and their parents. *Australian Journal of Education, 45*(3), 323-340.

Pech, S. (2013). Peer pals: supporting positive social-emotional behaviors in a kindergarten classroom. *Childhood Education, 89*(4), 217-223.

Porter, L. (2000). *Behaviour in schools: Theory and practice for teachers*. Buckingham, UK: Open University Press.

Powell, B. J., McMillen, J. C., Proctor, E. K., Carpenter, C. R., Griffey, R. T., Bunger, A. C., . . . York, J. L. (2012). A compilation of strategies for implementing clinical innovations in health and mental health. *Medical Care Research and Review, 69*(2), 123-157.

Rafferty, L. A. (2012). Self-monitoring during whole group reading instruction: Effects among students with emotional and behavioral disabilities during summer school intervention sessions. *Emotional and Behavioural Difficulties, 17*(2), 157-173.

Ramirez, E., Kulinna, P., & Cothran, D. (2012). Constructs of physical activity behaviour in children: The usefulness of social cognitive theory. *Psychology of Sport and Exercise, 13*(3), 303-310.

Rapp, J. T., Colby-Dirksen, A. M., Michalski, D. N., Carroll, R. A., & Lindenberg, A. M. (2008). Detecting changes in simulated events

using partial-interval recording and momentary time sampling. *Behavioral Interventions, 23*(4), 237-269.

Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 149-172). New York, NY: Springer.

Reeve, J., Deci, E. L., & Ryan, R. M. (2004). Self determination theory: A dialectical framework for understanding sociocultural influences on student motivation. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited* (Vol. 4, pp. 31-60). Greenwich, CT: Information Age Publishing.

Reiber, C., & McLaughlin, T. (2004). Classroom interventions: Methods to improve academic performance and classroom behavior for students with Attention-Deficit/Hyperactivity Disorder. *International Journal of Special Education, 19*(1), 1-13.

Reschly, A. L., & Christenson, S. L. (2012). Jingle, jangle, and conceptual haziness: Evolution and future directions of the engagement construct. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 3-19). New York, NY: Springer.

Robson, C. (2011). *Real world research* (3rd ed.). West Sussex, UK: John Wiley & Sons.

Rooney, K. J., Hallahan, D. P., & Lloyd, J. W. (1984). Self-recording of attention by learning disabled students in the regular classroom. *Journal of Learning Disabilities, 17*(6), 360-364.

Roth, A., & Fonagy, P. (2005). *What works for whom? A critical review of psychotherapy research* (2nd ed.). New York: The Guilford Press.

Rowe, K. J., & Rowe, K. S. (1992). The relationship between inattentiveness in the classroom and reading achievement (Part B): An explanatory study. *Journal of the American Academy of Child & Adolescent Psychiatry, 31*(2), 357-368.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78.

Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan

(Eds.), *Handbook of self-determination research* (pp. 3-34). Rochester, NY: The University of Rochester Press.

Salend, S. J., & Lamb, E. A. (1986). Effectiveness of a group-managed interdependent contingency system. *Learning Disability Quarterly, 9*(4), 268-273.

Salend, S. J., Reeder, E., Katz, N., & Russell, T. (1992). The effects of a dependent group evaluation system. *Education and Treatment of Children, 15*(1), 32-42.

Schanding Jr, G. T., & Sterling-Turner, H. E. (2010). Use of the mystery motivator for a high school class. *Journal of Applied School Psychology, 26*(1), 38-53.

Schultz, G. F., & Switzky, H. N. (1993). The academic achievement of elementary and junior high school students with behavior disorders and their nonhandicapped peers as a function of motivational orientation. *Learning and Individual Differences, 5*(1), 31-42.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage learning.

Shapiro, E. S. (2013). *Behavioral Observation of Students in Schools (BOSS)*. Bloomington, MN: Pearson.

Sheng, Y. (2008). Testing the assumptions of analysis of variance. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 324-340). Thousand Oaks, CA: Sage Publications.

Shih, S.-S. (2008). The relation of self-determination and achievement goals to Taiwanese eighth graders' behavioral and emotional engagement in schoolwork. *The Elementary School Journal, 108*(4), 313-334.

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31*(3), 351-380.

Skinner, B. F. (1938). *The behavior of organisms: An experimental analysis*. New York: D. Appleton-Century Company, Inc.

Skinner, B. F. (1953). *Science and human behavior*. New York: Simon & Schuster.

Skinner, C. H. (2004). Group contingencies. In T. S. Watson & C. H. Skinner (Eds.), *Encyclopedia of school psychology* (pp. 149-151). New York: Kluwer Academic / Plenum Publishers.

Skinner, C. H., Williams, R. L., & Neddenriep, C. E. (2004). Using interdependent group-oriented reinforcement to enhance academic performance in general education classrooms. *School Psychology Review, 33*(3), 384-397.

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21-44). New York, NY: Springer.

Smith, M. A., & Misra, A. (1992). A comprehensive management system for students in regular classrooms. *The Elementary School Journal, 92*(3), 353-371.

Solomon, R. W., & Wahler, R. G. (1973). Peer reinforcement control of classroom problem behavior. *Journal of Applied Behavior Analysis, 6*(1), 49-56.

Sparks, C., Dimmock, J., Whipp, P., Lonsdale, C., & Jackson, B. (2015). "Getting connected": High school physical education teacher behaviors that facilitate students' relatedness support perceptions. *Sport, Exercise, and Performance Psychology, 4*(3), 219-236.

Stage, S. A., & Quiroz, D. R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review, 26*(3), 333-368.

Standage, M., Duda, J. L., & Ntoumanis, N. (2005). A test of self-determination theory in school physical education. *British Journal of Educational Psychology, 75*(3), 411-433.

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences* (5th ed.). New York, NY: Routledge.

Stone, L. L., Otten, R., Engels, R. C., Vermulst, A. A., & Janssens, J. M. (2010). Psychometric properties of the parent and teacher versions of the strengths and difficulties questionnaire for 4-to 12-year-olds: A review. *Clinical Child and Family Psychology Review, 13*(3), 254-274.

Stroet, K., Opdenakker, M.-C., & Minnaert, A. (2013). Effects of need supportive teaching on early adolescents' motivation and engagement: A review of the literature. *Educational Research Review, 9*(1), 65-87.

Stronge, J. H., Tucker, P. D., & Hindman, J. L. (2004). *Handbook for qualities of effective teachers*. Alexandria, VA: Association for Supervision and Curriculum Development.

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the p value is not enough. *Journal of Graduate Medical Education, 4*(3), 279-282.

Taylor, G., Jungert, T., Mageau, G. A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: The unique role of intrinsic motivation. *Contemporary Educational Psychology, 39*(4), 342-358.

Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, CA: Sage Publications Inc.

Terenzi, C. M., Ervin, R. A., & Hoff, K. E. (2010). Implementation guidelines: Classwide self-management of rule following. *Journal of Evidence-Based Practices for Schools, 11*(2), 117-122.

Theodore, L. A., Bray, M. A., & Kehle, T. J. (2004). A comparative study of group contingencies and randomized reinforcers to reduce disruptive classroom behavior. *School Psychology Quarterly, 19*(3), 253-271.

Theodore, L. A., Bray, M. A., Kehle, T. J., & Jenson, W. R. (2001). Randomization of group contingencies and reinforcers to reduce classroom disruptive behavior. *Journal of School Psychology, 39*(3), 267-277.

Thomas, G. (2013). *How to do your research project: A guide for students in education and applied social sciences* (2nd ed.). London: Sage Publications Ltd.

Thorne, S., & Kamps, D. (2008). The effects of a group contingency intervention on academic engagement and problem behavior of at-risk students. *Behavior Analysis in Practice, 1*(2), 12-18.

Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The good behavior game: 1969-2002. *Behavior Modification, 30*(2), 225-253.

Traxson, D. (1994). Helping children to become more self-directing in their behaviour. In P. Gray, A. Miller, & J. Noakes (Eds.), *Challenging behaviour in schools: Teacher support, practical techniques and policy development* (pp. 223-240). London, UK: Routledge.

Trevino-Maack, S. I., Kamps, D., & Wills, H. (2015). A group contingency plus self-management intervention targeting at-risk secondary students' class-work and active engagement. *Remedial and Special Education, 36*(6), 347-360.

Trevithick, P. (2011). *Social work skills and knowledge: A practice handbook*. Berkshire, UK: Open University Press.

Trickey, S., & Topping, K. J. (2004). 'Philosophy for children': A systematic review. *Research Papers in Education, 19*(3), 365-380.

UK Local Area. (2016). UK Local Area. Retrieved from http://www.uklocalarea.com

University of Nottingham. (2013). *Code of research conduct and research ethics*. Nottingham, UK: University of Nottingham.

Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology, 44*(5), 331-349.

Urhahne, D. (2015). Teacher behavior as a mediator of the relationship between teacher judgment and students' motivation and emotion. *Teaching and Teacher Education, 45*, 73-82.

Van den Berghe, L., Cardon, G., Tallir, I., Kirk, D., & Haerens, L. (2016). Dynamics of need-supportive and need-thwarting teaching behavior: The bidirectional relationship with student engagement and disengagement in the beginning of a lesson. *Physical Education and Sport Pedagogy, 21*(6), 653-670.

Vance, M. J., Gresham, F. M., & Dart, E. H. (2012). Relative effectiveness of DRO and self-monitoring in a general education classroom. *Journal of Applied School Psychology, 28*(1), 89-109.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The kappa statistic. *Family Medicine, 37*(5), 360-363.

Ward, J. (1971). Modification of deviant classroom behaviour. *British Journal of Educational Psychology, 41*(3), 304-313.

Webber, J., Scheuermann, B., McCall, C., & Coleman, M. (1993). Research on self-monitoring as a behavior management technique in special education classrooms a descriptive review. *Remedial and Special Education, 14*(2), 38-56.

Weeden, M., Wills, H. P., Kottwitz, E., & Kamps, D. (2016). The effects of a class-wide behavior intervention for students with emotional and behavioral disorders. *Behavioral Disorders, 42*(1), 285-293.

Wehlitz, R. C. (1960). Pupil-managed classes: A lesson in self-discipline. *The Clearing House, 34*(7), 393-397.

Weiner, I. B., & Craighead, W. E. (2010). *The Corsini encyclopedia of psychology* (4th ed. Vol. 3). Hoboken, NJ: John Wiley & Sons.

Wentzel, K. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology, 89*(3), 411-419.

Wentzel, K. (1998). Social relationships and motivation in middle school: The role of parents, teachers, and peers. *Journal of Educational Psychology, 90*(2), 202-209.

Wheldall, K., & Merrett, F. (1984). The behavioural approach to classroom management. In D. Fontana (Ed.), *Behaviourism and learning theory in education* (pp. 15-41). Edinburgh, UK: Scottish Academic Press Limited.

Williamson, B. D., Campbell-Whatley, G. D., & Lo, Y. Y. (2009). Using a random dependent group contingency to increase on-task behaviors of high school students with high incidence disabilities. *Psychology in the Schools, 46*(10), 1074-1083.

Willower, D. J. (1960). Leadership styles and leaders' perceptions of subordinates. *Journal of Educational Sociology, 34*(2), 58-64.

Wills, H., Kamps, D., Fleming, K., & Hansen, B. (2016). Student and teacher outcomes of the class-wide function-related intervention team efficacy trial. *Exceptional Children, 83*(1), 58-76.

Wills, H. P., Iwaszuk, W. M., Kamps, D., & Shumate, E. (2014). CW-FIT: Group contingency effects across the day. *Education and Treatment of Children, 37*(2), 191-210.

Wolfe, L. H., Heron, T. E., & Goddard, Y. L. (2000). Effects of self-monitoring on the on-task behavior and written language performance of elementary students with learning disabilities. *Journal of Behavioral Education, 10*(1), 49-73.

Wright, C., Weekes, D., & McGlaughlin, A. (2000). *Race, class and gender in exclusion from school*. Oxon, UK: Routledge.

Wright, R. A., & McCurdy, B. L. (2012). Class-wide positive behavior support and group contingencies: Examining a positive variation of the good behavior game. *Journal of Positive Behavior Interventions, 14*(3), 173-180.

Wubbels, T. (2011). An international perspective on classroom management: what should prospective teachers learn? *Teaching Education, 22*(2), 113-131.

Yawkey, T. D. (1971). Conditioning independent work behavior in reading with seven-year-old children in a regular early childhood classroom. *Child Study Journal, 2*(1), 23-34.

Yeo, L. S., & Choi, P. M. (2011). Cognitive-behavioural therapy for children with behavioural difficulties in the Singapore mainstream school setting. *School Psychology International, 32*(6), 616-631.

Yoder, P., & Symons, F. (2010). *Observational measurement of behavior*. New York, NY: Springer Publishing Company.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA: Academic Press.

8. Appendices

8.1. **Appendix: Search strategy**

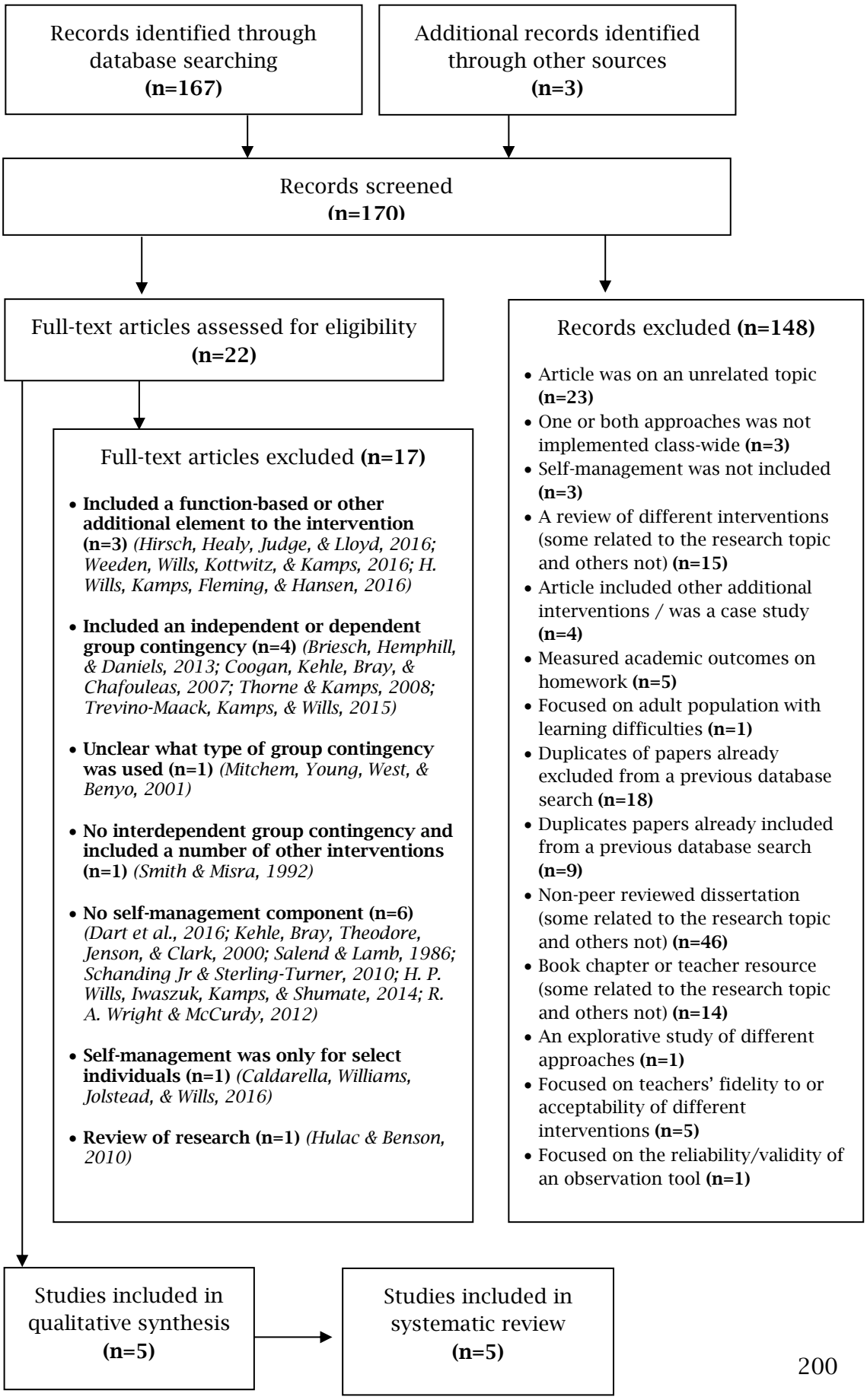
Database searched	**Search terms used**	**Papers found**	**Number included**	**Number excluded**	**Studies included**	**Date searched**	**Search dates**
Web of Science	Self-management AND group contingency	20	2	18	(Davies & Witte, 2000; Hoff & Ervin, 2013)	23rd March 2017	1900-2017
PsycINFO (Ovid)	Self-management AND group contingency	10	1	9	(Chafouleas et al., 2012)	23rd March 2017	1806-2017
ERIC (EBSCO)	Self-management AND group contingency	14	2	12	(Coogan et al., 2007; Denune et al., 2015)	23rd March 2017	1986-2017
IngentaConnect	Self-management AND group contingency	3	0	3	N/A	23rd March 2017	1998-2017
Google Scholar	“self-	120	0	120		23rd March	Any time

	management" AND "interdependent group contingency" AND "class"					2017	

Table 8.1: A table to outline the databases and search terms used to locate journal articles in the systematic review, as well as the numbers of papers included and excluded.

See appendix 8.2 for information on why some of the articles were excluded.

8.2. **Appendix: Screening flow diagram**



8.3. **Appendix: Study characteristics**

Study	Sample	School Demographic characteristics	Study design & Conditions	Dependent variable	Location	Length of intervention	Outcome measures
Chafouleas, Sanetti, Jaffery & Fallon (2012)	12-13 year olds 2 teachers, 57 pupils 81% White; 13% Hispanic; 2% Asian American; 6% Biracial	Mainstream middle school 20% low income pupils	Multiple baseline design with embedded changing criterion Self-management THEN interdependent group contingency added	Preparedness, engagement (on- and off-task), homework completion	US	Variable between 4 – 6 weeks	Direct Behavior Rating form *(for pupils to measure their own behaviour)* Systematic Direct Observation form *(for researchers to measure engagement)* 15 min observations using 15 sec interval momentary time sampling *(for on-task – one pupil observed at a time, at random)* and partial-interval recording *(for off-task behaviour)* Usage Rating Profile Intervention *(self-report tool for teachers to determine intervention usability)*

Study	Sample	School Demographic characteristics	Study design & Conditions	Dependent variable	Location	Length of intervention	Outcome measures
Davies & Witte (2000)	7-8 year olds 1 teacher, 30 pupils Data collected on ADHD participants, with matched controls (n = 8) 100% white participants, 8-10 years old	Mainstream elementary school	ABAB reversal design Self-management AND group contingency together	Inappropriate verbalisations	US	22 days	30 min observations Event recording *(for frequency of inappropriate verbalisations)*

Study	Sample	School Demographic characteristics	Study design & Conditions	Dependent variable	Location	Length of intervention	Outcome measures
Denune, Hawkins, Donovan, McCoy, Hall & Moeder (2015)	10-11 year olds 1 teacher, 16 pupils in the class 12-15 years old Data collected on 14 pupils: White pupils (n = 4); Black pupils (n = 10) ADHD, ODD and PTSD diagnoses	Alternative school for pupils with 'emotional and behavioural disorders'	ABCBC Withdrawal design Group contingency THEN self-management added	On-task, off-task and disruptive behaviours	US	Intervention conditions varied between 6-9 days	40-45 min daily observations Adapted Behavioral Observation of Students in Schools, to measure engagement and disruption On- and off- task behaviours measured using 20 sec momentary interval recording (one pupil observed at a time in fixed order) Disruptive behaviours measured using 20 sec partial interval recording (data tallied by frequency) Social validity questionnaire *(for teachers to determine intervention usability)*

Study	Sample	School Demographic characteristics	Study design & Conditions	Dependent variable	Location	Length of intervention	Outcome measures
Glynn, Thomas & Shee (1973)	6-7 year olds 1 teacher, 37 pupils Pupils 6-7 years old Data collected on whole class and n = 8 randomly picked pupils	No information given	ABCACDEEAE design Group contingency only THEN self-management only	On-task behaviour	New Zealand	85 total days of the study	30 minute observations 10 sec interval recording of individuals' and whole group on-task behaviour

Study	Sample	School Demographic characteristics	Study design & Conditions	Dependent variable	Location	Length of intervention	Outcome measures
Hoff & Ervin (2013)	6-7 year olds 3 teachers, 64 pupils Data collected on whole class and n = 3 "at risk" pupils (2 with ADHD)	Mainstream elementary school 82.7% Caucasian population 8.7% low income pupils	Multiple baseline across subjects design Group contingency THEN Self-management added	Decreasing disruptive behaviours	US	Variable between 8 – 26 sessions	35 minute observations 15 sec partial-interval recording of disruptive behaviours Adapted Intervention Rating Profile and Children's Intervention Rating Profile *(to determine intervention usability)*

Table 8.2: A table to summaries the characteristics of the studies reviewed in the systematic review.

8.4. **Appendix: Weight of evidence ratings**

Studies	Weight of Evidence A												
	Sample size	Intervention described sufficiently	Researcher affiliation with treatment / stakeholder	Use of a control / comparison group	Intervention fidelity monitored / had been trained	Pre- and post- (P)/ repeated (R) measures	Follow-up measures	Drop-out rate	Standardised measurement tools	Blinding of outcome assessors	Clear Confounding variables	Equivalent Groups	Inter-observer agreement
Chafouleas, Sanetti, Jaffery & Fallon (2012)	57	✓	×	×	✓	R	×	0	×	×	×	×	92.4%
Davies & Witte (2000)	8	✓	×	×	✓	R	×	0	×	×	✓	N/A	82-87%
Denune, Hawkins, Donovan, McCoy, Hall & Moeder (2015)	14	✓	×	×	✓	R	✓	0	×	×	×	N/A	96.3%
Glynn, Thomas & Shee (1973)	37	✓	×	×	×	R	×	0	×	×	✓	N/A	84-90%+
Hoff & Ervin (2013)	84	✓	×	×	✓	R	×	0	×	✓	×	×	92.5%

Table 8.3: A table to show the Weight of Evidence A judgements for the systematic review papers.

Studies Included	Weight of Evidence A	Weight of Evidence B	Weight of Evidence C	Weight of Evidence D
	The coherence and integrity of the evidence in its own terms.	Appropriateness of this form of evidence for review question	Appropriateness of the focus of the research, for answering the review question	Overall judgement of quality and contribution towards answering the review question
Chafouleas, Sanetti, Jaffery & Fallon (2012)	**Medium**	**Medium** *(Did not measure the impact of self-management alone)*	**High**	**Medium**
Davies & Witte (2000)	**Medium**	**Medium** *(only measured one type of disruptive behaviour for 8 pupils only)*	**Low** *(only measured one type of disruptive behaviour for 8 pupils only)*	**Medium**
Denune, Hawkins, Donovan, McCoy, Hall & Moeder (2015)	**High**	**High**	**Medium** *(Alternative school population)*	**High**

Studies Included	Weight of Evidence A	Weight of Evidence B	Weight of Evidence C	Weight of Evidence D
Glynn, Thomas & Shee (1973)	**Low** *(Intervention was not tightly controlled. Teacher changed aspects. Difficulties with measurement by student observers)*	**Low** *(Findings may be invalid or unreliable due to lack of control and gaps in data collection)*	**High** *(Looked at group contingency separate from self-management with the whole class)*	**Low**
Hoff & Ervin (2013)	**Medium**	**High**	**High**	**High**

Table 8.4: Weight of Evidence ratings for the reviewed papers.

8.5. **Appendix: Ethical approval letters**

School of Psychology
The University of Nottingham
University Park
Nottingham
NG7 2RD

tel: +44 (0)115 846 7403 or (0)115 951 4344

SJ/wb
Ref: 820

Wednesday, 20 April 2016

Dear Kamal Bhana & Nathan Lambert,

Ethics Committee Review

Thank you for submitting an account of your proposed research 'The Impact of Self-Management and Interdependent Group Contingency Approaches on Whole-class Behaviour'.

That proposal has now been reviewed and we are pleased to tell you it has met with the Committee's approval.

However:

Please note the following comments from our reviewers;

- The statement "which has been approved by the University of Nottingham Ethics Committee." in the letter to the head teacher should be corrected. It is not the University of Nottingham Ethics Committee but the School of Psychology Ethics Committee.
- The parent information sheet should be written using less specialised language (e.g., "inter-dependent group contingency approach").
- Parent consent form: "...except observation data that has already been conducted)" not sure this needs to be included in the consent form because this data is not linked to a specific pupil that can be identified? If it is the data could be removed.

Final responsibility for ethical conduct of your research rests with you or your supervisor. The Codes of Practice setting out these responsibilities have been published by the British Psychological Society and the University Research Ethics Committee. If you have any concerns whatever during the conduct of your research then you should consult those Codes of Practice. The Committee should be informed immediately should any participant complaints or adverse events arise during the study.

Independently of the Ethics Committee procedures, supervisors also have responsibilities for the risk assessment of projects as detailed in the safety pages of the University web site. Ethics Committee approval does not alter, replace, or remove those responsibilities, nor does it certify that they have been met.

Yours sincerely

Professor Stephen Jackson
Chair, Ethics Committee

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

School of Psychology
The University of Nottingham
University Park
Nottingham
NG7 2RD

tel: +44 (0)115 846 7403 or (0)115 951 4344

SJ/wb
Ref: 894

Monday, 10 October 2016

Dear Kamal Bhana & Nathan Lambert,

Title of the new project: **The Impact of Self-Management and Inter-dependent Group Contingency Approaches on Whole-class Behaviour**
Applicants: Kamal Bhana & Nathan Lambert

Further to your request for Chair Approval for amendments to the project:-

Details of the previous study:

Applicant: Kamal Bhana
Title: Trainee Educational Psychologist
Date of approval: 20.04.2016
Reference number (if known): 820

As Chair of the Ethics Committee I have considered your request and I am happy to grant approval for the following changes:

Change to consent procedure i.e. moving from 'opt-in' to 'opt-out' consent for the data collection/ data sharing aspect

Final responsibility for ethical conduct of your research rests with you or your supervisor. The Codes of Practice setting out these responsibilities have been published by the British Psychological Society and the University Research Ethics Committee. If you have any concerns whatever during the conduct of your research then you should consult those Codes of Practice.

Independently of the Ethics Committee procedures, supervisors also have responsibilities for the risk assessment of projects as detailed in the safety pages of the University web site. Ethics Committee approval does not alter, replace, or remove those responsibilities, nor does it certify that they have been met.

Yours sincerely

Professor Stephen Jackson
Chair, Ethics Committee

8.6. **Appendix: Stakeholder engagement letters**

Educational Psychology Service
[redacted]
Tel: [redacted]

[redacted]

20th May 2016

Dear [redacted],

The Impact of Self-Management and Inter-dependent Group Contingency Approaches on Whole-class Behaviour

My name is Kamal Bhana. I am currently undertaking a Doctorate in Applied Educational Psychology at the University of Nottingham and I am on placement at [redacted] Educational Psychology Service. For my thesis, I am planning to undertake a research project to evaluate the effectiveness of two whole-class approaches for increasing on-task behaviour and reducing disruptive behaviour in general key stage 2 classrooms. I am writing to ask whether you would be interested in your school being involved in this research project, which has been approved by the School of Psychology Ethics Committee.

I hope that the findings from this research project will indicate that one or both of these approaches are successful in increasing engagement. There is some evidence in the research literature which suggests that these approaches have been effective in other settings.

I would wish to work alongside three teachers teaching in the same year group. This would preferably be Year 4 or Year 5 but not necessarily. These teachers must be interested in volunteering their participation. I would expect to be working alongside these members of staff for around 14 weeks. The project will involve two of the teachers implementing both approaches and the third will receive no approach at this time, in order to be the waitlist control group. The class that receives no approach during this study, will be provided with access to the approach(es) at the end of the study, if one or more were found to be effective. Before classes can be recruited for the study, I would need to do some observations to see whether the approaches would be suitable for those classes.

The teachers will fill out a short questionnaire for each pupil in the class, three times throughout the course of the study. I will also conduct nine classroom observations in each class during the course of the study, to measure levels of whole-class engagement. I expect that I would need to meet with the teachers prior to starting the study and half way through, to train them on delivering the approaches.

I can assure confidentiality with the data acquired. Nobody other than my University tutor, placement supervisor, external examiners and me will have access to the data. The identities of the school, teacher and children will also be kept confidential and will be unidentifiable by outside persons. The limits of confidentiality of course are that staff members within the school may know which teachers are participating in the study. Furthermore, parents of the children in the class will be aware, as their consent would be sought for data to be collected on their children. Following completion of the study, the teacher, pupils and their parents/carers will be debriefed about the aims of the study and the findings overall.

If you are interested in your school participating, I would encourage you to approach your teachers to gage the level of interest. If you choose to proceed, please contact me via email at *lpxknb@nottingham.ac.uk* to organise a meeting, where we can discuss this further. I hope that this research project will be of great benefit to your school.

Please understand that should you choose to proceed, your school's participation is entirely voluntary. The school maintains the right to withdraw at any stage of the study, including up to six weeks after the completion of the study.

Thank you for taking the time to read this letter and I look forward to hearing from you.

Yours sincerely,

Kamal Bhana
Trainee Educational Psychologist

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

8.7. Appendix: Head teacher and Chair of Governers' research information sheet and consent form

School of Psychology
Head Teacher / Chair of Governors Information Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: **The Impact of Self-Management and Interdependent Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number: *820*
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

This is an invitation to take part in a research study on the effectiveness of two whole-class approaches for improving behaviour in the classroom. These approaches are self-management and interdependent group contingency.

Before you decide if you wish to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

I am a Trainee Educational Psychologist at the University of Nottingham. As part of my Doctoral research project, I am interested in seeing whether one of these approaches that can be implemented at a whole-class level, is more effective than the other in increasing overall on-task behaviours in the classroom.

WHAT WILL HAPPEN:

If you participate, the Year 4 classes in your school will either receive self-management first and interdependent group contingency after a few weeks, or interdependent group contingency first and self-management after a few weeks, or neither (waitlist control). The classes in the waitlist control group will receive the approach after the study has ended, if one or more of the approaches is shown to be effective.

In the classes that receive self-management, all that the children will be doing differently to usual, is they will be asked to think about the class rules and rate themselves on how well they were following that rule. If they are in the classes receiving group contingency, they will participate in rating the whole class on how well classroom rules were followed. They will not be required to participate in any special tests. They will simply be observed in their usual lessons. Children who are not receiving

the approach at this time will not be expected to do anything at all that is different from the usual.

In order to measure the effectiveness of these approaches, the teachers will be asked to fill out a brief questionnaire for each pupil in the class at three different time points during the study. This questionnaire measures the child's strengths and difficulties within a lesson. I shall also conduct nine classroom observations in which I collect data on on- and off-task behaviours. These observations will last between 30 and 45 minutes each. Finally, both approaches require the children to make a judgement on whether they followed the class rules, by rating themselves and the class on a scale of 0 to 4. I will also need to collect this data for the research.

For those teachers receiving these approaches, they will receive training on how to do the approach and soon after, I will observe them putting it into action to support them in implementing the procedures correctly.

TIME COMMITMENT

It is expected that once the approach is up and running in the classroom, all of the data collection should be completed within 14 weeks.

It is estimated that filling in the questionnaire for each pupil should take around 2 hours in total to complete. This will only need to be done three times over the course of the 14 week study. Each time, the teachers will be given a week in which to complete it.

BENEFITS AND RISKS

This study poses no known risks to teachers or the children. It is hoped that the approaches will benefit pupils in helping them to concentrate on their learning, thereby being of benefit to teachers as well. The study will also help to deepen our understanding of whether these approaches are effective in improving engagement in lessons.

CONFIDENTIALITY/ANONYMITY

The data collected will remain confidential. Nobody other than my University tutor and I will have access to the questionnaires that are filled out on each pupil. This information will not be available to anybody who may ask for it. The observation data will only list numbers of pupils; no names. As such, anonymity of each pupil is guaranteed. When analysing the data, none of the children will be identifiable in any way and once analysis is complete, the questionnaires will be destroyed. The overall class data collected will be published in a thesis and may be presented at a conference but no individual children will be identifiable.

Participation in this study is totally voluntary and you are under no obligation to take part. You are free to withdraw at any point before or during the study. All data collected will be kept confidential and used for research purposes only. It will be stored in compliance with the Data Protection Act.

FOR FURTHER INFORMATION

I will be glad to answer your questions about this study at any time, and can inform you about the results of the study once data collection is complete.

You may contact me at the following email address: lpxknb@nottingham.ac.uk

Once you have had your queries answered, if you decide to participate in this study, you will be asked to sign the consent form below.

If you have any questions or concerns please don't hesitate to ask now. We can also be contacted after your participation at the above address.

Kamal Bhana
Trainee Educational Psychologist

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

School of Psychology
Head Teacher / Chair of Governors Information Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: **The Impact of Self-Management and Interdependent Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number: *820*
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

The participant should answer these questions independently:

- Have you read and understood the Information Sheet?
  YES/NO

- Have you had the opportunity to ask questions about the study?
  YES/NO

- Have all your questions been answered satisfactorily?
  YES/NO

Do you understand that you are free to withdraw from the study (at any time and without giving a reason)?
YES/NO

- I give permission for data from this study to be shared with other researchers provided that my anonymity is completely protected.
  YES/NO

- Do you agree to take part in the study?
  YES/NO

"This study has been explained to me to my satisfaction, and I agree to take part. I understand that I am free to withdraw at any time."

Signature of the Head Teacher: Date:

Name (in block capitals)

Signature of the Chair of Governors: Date:

Name (in block capitals)

I have explained the study to the above participant and he/she has agreed to take part.

Signature of researcher: Date:

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

8.8. Appendix: Teachers' research information sheet and consent form

School of Psychology
Teacher Information

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: **The Impact of Self-Management and Interdependent Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number: *820*
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

This is an invitation to take part in a research study on the effectiveness of two whole-class approaches for improving behaviour in the classroom. These approaches are self-management and interdependent group contingency.

Before you decide if you wish to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

I am a Trainee Educational Psychologist at the University of Nottingham. As part of my Doctoral research project, I am interested in seeing whether one of these approaches that can be implemented at a whole-class level, is more effective than the other in increasing overall on-task behaviours in the classroom.

WHAT WILL HAPPEN:

If you participate, your class will either receive self-management first and interdependent group contingency after a few weeks, or interdependent group contingency first and self-management after a few weeks, or neither (waitlist control). If your class is in the waitlist control group and therefore receives no approach at this time, rest assured that you will receive the approach after the study has ended, if one or more of the approaches is shown to be effective.

If your class receives self-management, all that the children will be doing differently to usual, is they will be asked to think about the class rules and rate themselves on how well they were following that rule. If they are in the class receiving group contingency, they will participate in rating the whole class on how well classroom rules were followed. They will not be required to participate in any special tests. They will simply be

observed in their usual lessons. Children who are not receiving the approach at this time will not be expected to do anything at all that is different from the usual.

In order to measure the effectiveness of these approaches, you will be asked to fill out a brief questionnaire for each pupil in your class at three different time points during the study. This questionnaire measures the child's strengths and difficulties within a lesson. I shall also conduct nine classroom observations in which I collect data on on-task, off-task and disruptive behaviours. These observations will last around 30 minutes each. Finally, both approaches require the children to make a judgement on whether they followed the class rules, by rating themselves and the class on a scale of 0 to 4. I will also need to collect this data for the research.

For those teachers receiving these approaches, you will receive training on how to do the approach and soon after, I will observe you putting it into action to support you in implementing the procedures correctly.

TIME COMMITMENT

It is expected that once the approach is up and running in the classroom, all of the data collection should be completed within 14 weeks.

It is estimated that filling in the questionnaire for each pupil should take around 2 hours in total to complete. This will only need to be done three times over the course of the 14 week study. Each time, you will be given a week in which to complete it.

BENEFITS AND RISKS

This study poses no known risks to you or the children. It is hoped that the approaches will benefit pupils in helping them to concentrate on their learning, thereby being of benefit to teachers as well. The study will also help to deepen our understanding of whether these approaches are effective in improving engagement in lessons.

CONFIDENTIALITY/ANONYMITY

The data collected will remain confidential. Nobody other than my University tutor and I will have access to the questionnaires that are filled out on each pupil. This information will not be available to anybody who may ask for it. The observation data will only list numbers of pupils; no names. As such, anonymity of each pupil is guaranteed. When analysing the data, none of the children will be identifiable in any way and once analysis is complete, the questionnaires will be destroyed. The overall

class data collected will be published in a thesis and may be presented at a conference but no individual children will be identifiable.

Participation in this study is totally voluntary and you are under no obligation to take part. You are free to withdraw at any point before or during the study. All data collected will be kept confidential and used for research purposes only. It will be stored in compliance with the Data Protection Act.

FOR FURTHER INFORMATION

I will be glad to answer your questions about this study at any time, and can inform you about the results of the study once data collection is complete.

You may contact me at the following email address: lpxknb@nottingham.ac.uk

Once you have had your queries answered, if you decide to participate in this study, you will be asked to sign the consent form below.

If you have any questions or concerns please don't hesitate to ask now. We can also be contacted after your participation at the above address.

Kamal Bhana
Trainee Educational Psychologist

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

School of Psychology
Teacher Consent Form

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: **The Impact of Self-Management and Interdependent Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number: *820*
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

The participant should answer these questions independently:

- Have you read and understood the Information Sheet?
  YES/NO
- Have you had the opportunity to ask questions about the study?
  YES/NO
- Have all your questions been answered satisfactorily?
  YES/NO
- Do you understand that you are free to withdraw from the study (at any time and without giving a reason)?
  YES/NO
- I give permission for my data from this study to be shared with other researchers provided that my anonymity is completely protected.
  YES/NO
- Do you agree to take part in the study?
  YES/NO

"This study has been explained to me to my satisfaction, and I agree to take part. I understand that I am free to withdraw at any time."

Signature of the Participant: Date:

Name (in block capitals)

I have explained the study to the above participant and he/she has agreed to take part.

Signature of researcher: Date:

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

8.9. Appendix: Parent information sheets and consent forms

School of Psychology
Parent Information Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: **The Impact of Self-Management and Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number: *820*
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

This is an invitation for your child to take part in a research study looking at how effective two approaches for classroom management are in helping children stay on-task in lessons. These approaches are self-management and group contingency.

Before you decide whether you wish for your child to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

WHAT WILL HAPPEN:

Your child's class will either receive both approaches or none of them at this time. If your child's class does not receive the approach at this time, they will receive an approach after the study has ended, if one or more of the approaches is shown to be effective.

The only thing that your child will be asked to do that is different from normal is:

- If your child's class receives self-management, they will be asked to think about the class rules and rate themselves on how well they were following that rule.
- If they receive group contingency, they will rate the class as a whole on how well they followed the classroom rules.

They will not be required to participate in any special tests. Children who are not receiving the approach at this time will not be expected to do anything at all that is different from the usual.

In order to find out if these approaches work, your child's class teacher will fill out questionnaires on each pupil in the class, at three different time points during the research. I will also observe each class six times in total, to measure the overall levels of engagement of the whole class. I will not be focusing on individual children. These observations will last around 30-45 minutes each. During the lessons, the children will be asked to rate themselves and their class on how well they followed the rules. This information will be collected as well, for the research and reported in the final write-up but it will be reported anonymously and no children will be identifiable.

TIME COMMITMENT

It is expected that once the approach is up and running in the classroom, all of the data collection should be completed within 14 weeks.

BENEFITS AND RISKS

This study poses no known risks to the teacher or the children. It is hoped that the approaches will benefit pupils in helping them to concentrate on their learning. The study will also help to deepen our understanding of whether these approaches are effective in improving on-task behaviour in lessons.

CONFIDENTIALITY/ANONYMITY

The data collected will remain confidential. Nobody other than my University tutor, examiner and I will have access to the questionnaires that are filled out on each pupil. This information will not be available to anybody who may ask for it. The observation data will only produce an overall score for the whole class; no names will be taken. As such, anonymity of each pupil is guaranteed. When analysing the data, none of the children will be identifiable in any way and once analysis is complete, the questionnaires will be destroyed. The overall class data collected will be published in a thesis, published online and may be presented at a conference but no individual children will be identifiable.

Participation in this study is totally voluntary and your child is under no obligation to take part. Although your child cannot be removed from receiving the approach as it is a whole-class approach, you are free to withdraw your child's individual data at any point before or during the study, and up to six weeks after the end of the study. All data collected will be kept confidential and used for research purposes only. It will be stored in compliance with the Data Protection Act and destroyed at the end of the research.

FOR FURTHER INFORMATION

I will be holding a parents' meeting at your child's school on **Monday 26th September** at **9.00am** and **2.30pm**, in order to answer any questions that you might have about this study. Please come along. If you cannot make this date, I will be glad to answer your questions at any time, and can inform you about the results of the study once data collection is complete.

If you have any questions or concerns please don't hesitate to ask now. We can also be contacted after your participation at the above address.

Please return the attached consent form to the school as soon as possible so that I may know whether or not you wish for your child to take part. **Returning the form will automatically enter your child into a PRIZE DRAW, even if consent is not given to take part in the research.** The draw will take place in October and three children will be presented with **gift vouchers for either £20 (first prize), £15 (second prize) or £10 (third prize).** If the form is not returned completed, they will not be entered.

Kamal Bhana
Trainee Educational Psychologist

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

School of Psychology
Parent Consent Form

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: **The Impact of Self-Management and Interdependent Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number: *820*
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

The participant should answer these questions independently:

- Have you read and understood the Information Sheet?
  YES/NO
- Have you had the opportunity to ask questions about the study?
  YES/NO
- Have all your questions been answered satisfactorily?
  YES/NO
- Do you understand that you are free to withdraw your child's data from the study at any time and without giving a reason?
  YES/NO
- I give permission for my child's data from this study to be shared with other researchers provided that my child's anonymity is completely protected.
  YES/NO
- Do you agree to allow your child to take part in the study?
  YES/NO

"This study has been explained to me to my satisfaction, and I agree to allow my child to take part. I understand that I am free to withdraw their data at any time."

Signature of the Parent/carer: Date:

Name of the child (in block capitals):

I have explained the study to the above participant and he/she has agreed to take part.

Signature of researcher: Date:

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email:***
nathan.lambert@nottingham.ac.uk
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Dear Parent/Carer,

Change in research, in Year 4!

Thank you to all parents who returned the consent form for the research. Your child's name has been entered into the prize draw. The draw will take place next week.

There has been a slight change in the research design, which means that your child's information **will be included in the research** automatically, **unless you request for your child's data to not be included.** I have attached the amended parent information sheet for your reference.

If you <u>are happy</u> for your child's information to be included in the research, you do NOT need to do anything.

If you want your child's data <u>to be removed</u>, please fill out the 'Opt-out' consent form attached and return it to your school by **Friday 14th October**.

If you have any questions, feel free to email me at lpxknb@nottingham.ac.uk. I will also be holding a parents' meeting at **3.15pm** on **Wednesday 12th October** at the school to answer any questions that you might have.

Thank you for your time,

Kamal Bhana
(Trainee Educational Psychologist)

School of Psychology
Parent Information Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: **The Impact of Self-Management and Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number: *820*
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

This is an invitation for your child to take part in a research study looking at how effective two approaches for classroom management are in helping children stay on-task in lessons. These approaches are self-management and group contingency.

Before you decide whether you wish for your child to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

I am a Trainee Educational Psychologist at the University of Nottingham. As part of my Doctoral research project, I am interested in seeing whether one of these two approaches is more effective than the other in improving on-task behaviours in the classroom.

WHAT WILL HAPPEN:

Your child's class will either receive both approaches or none of them at this time. If your child's class does not receive the approach at this time, they will receive an approach after the study has ended, if one or more of the approaches is shown to be effective.

The only thing that your child will be asked to do that is different from normal is:

- If your child's class receives self-management, they will be asked to think about the class rules and rate themselves on how well they were following that rule.
- If they receive group contingency, they will rate the class as a whole on how well they followed the classroom rules.

They will not be required to participate in any special tests. Children who are not receiving the approach at this time will not be expected to do anything at all that is different from the usual.

In order to find out if these approaches work, your child's class teacher will fill out questionnaires on each pupil in the class, at three different time points during the research. I will also observe each class nine times in total, to measure the overall levels of engagement of the whole class. I will not be focusing on individual children. These observations will last around 30 minutes each. During the lessons, the children will be asked to rate themselves and their class on how well they followed the rules. This information will be collected as well, for the research and reported in the final write-up but it will be reported anonymously and no children will be identifiable.

Before the teachers pass the questionnaires they have completed and the ratings provided by the children themselves to me, they will apply a code so that I am not able to match the information provided to individual children. The information will not be shared with any other parties.

In future reports the information will only be reported anonymously. No children will be identifiable.

TIME COMMITMENT

It is expected that once the approach is up and running in the classroom, all of the data collection should be completed within 14 weeks.

BENEFITS AND RISKS

This study poses no known risks to the teacher or the children. It is hoped that the approaches will benefit pupils in helping them to concentrate on their learning. The study will also help to deepen our understanding of whether these approaches are effective in improving on-task behaviour in lessons.

CONFIDENTIALITY/ANONYMITY

The data collected will remain confidential. Nobody other than my University tutor, examiner and I will have access to the questionnaires that are filled out on each pupil – and data will be labelled by code prior to being passed to me. This information will not be available to anybody who may ask for it. The observation data will only produce an overall score for the whole class; no names will be taken. As such, anonymity of each pupil is guaranteed. When analysing the data, none of the children will be identifiable in any way and once analysis is complete, the questionnaires will be destroyed. The overall class data collected will be

published in a thesis, published online and may be presented at a conference but no individual children will be identifiable.

Participation in this study is totally voluntary and your child is under no obligation to take part. Although your child cannot be removed from receiving the approach as it is a whole-class approach, you are free to withdraw your child's individual data at any point before or during the study, and up to six weeks after the end of the study.

If you wish to withdraw your child's data at any point, you should complete the attached 'opt-out' consent form and return it, by **Friday 14th October**, to:

Kamal Bhana email: lpxknb@nottingham.ac.uk
Or
Your child's class teacher

All data collected will be kept confidential and used for research purposes only. It will be stored in compliance with the Data Protection Act and destroyed at the end of the research.

FOR FURTHER INFORMATION

I will be holding a parents' meeting at your child's school on **Wednesday 12th October at 3.15pm**, in order to answer any questions that you might have about this study. If you cannot make this date, I will be glad to answer your questions any time, and can inform you about the results of the study once data collection is complete.

If you have any questions or concerns please don't hesitate to ask now. I can also be contacted after your participation at the above address.

Kamal Bhana
Trainee Educational Psychologist

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email:***
nathan.lambert@nottingham.ac.uk
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

School of Psychology
Parent Consent Form

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Opt-out Consent Form for parent/carer

Title of Project: **The Impact of Self-Management and Inter-dependent Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number or Taught Project Archive Number:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

"This study has been explained to me to my satisfaction, and I prefer NOT to take part."

Signature of the Participant (pupil): _____ Date _____

Name (in block capitals): _____

Signature of the Parent/Carer: _____ Date:_____

Name (in block capitals): _____

I have explained the study to the above participant and he/she wishes to NOT take part.

Signature of researcher: _____ Date:_____

If you have any questions or complaints about the study, please contact:
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***
or
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

8.10. **Appendix: Approach procedures**

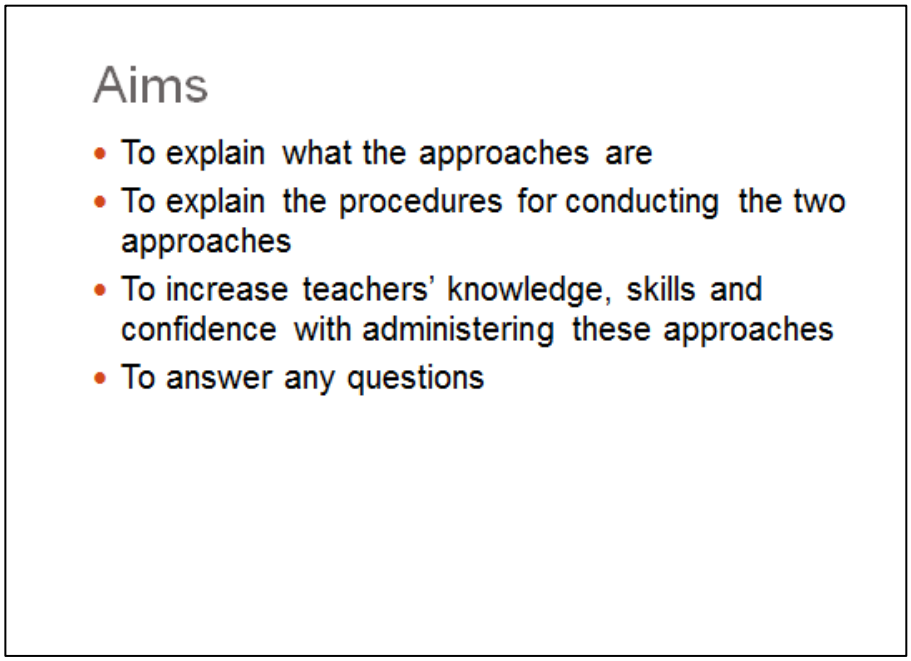
Conditions	Details
Self-management only	Teacher trains the pupils on how to self-rate accurately. During the training, the teacher outlines/reminds the pupils of the classroom rules and the class discuss examples of rule-breaking and rule-following behaviours. Following this, self-management is implemented in one lesson every afternoon. The teacher begins each lesson by reminding the class of the rules. Every 15 minutes, the teacher chooses one rule at random and the pupils rate on a scale, to what degree they were following that rule in the 15 minutes prior: 0= not at all; 1= some of the time; 2= half of the time; 3= most of the time; 4= all of the time. At the end of the lesson, the pupils rate the whole class on a rule randomly chosen by the teacher, on the 0-4 scale. The majority vote is taken as the class vote and the teacher shares his/her rating for the whole class. At the end of the lesson, the pupils spend a few minutes sharing their ratings with their partner and justifying their choice. Partners give feedback to each other on whether they agree with their partner's ratings.

Interdependent group contingency only	The teacher begins by training the pupils. The class rules are outlined and discussed. The contingency for earning group rewards is shared. Ground rules are set to inform pupils to be supportive and respectful of each other, and not to say upsetting things to a peer who may not have followed a rule. At the start of each lesson, a quick reminder of the rules is provided. During the lesson, the teacher keeps track of rule following and breaking behaviour of the whole class. At the end of the lesson, the teacher chooses at random, one of the classroom rules and awards the class between 0 and 4 points depending on how well they as a class followed the rule. The points are marked on a graph which is clearly displayed. This is done in front of the pupils and will lead to a reward once the pre-determined criterion is reached. The reward would be negotiated between the teacher and pupils at the start of this intervention phase.
Interdependent group contingency & self-management	The teacher trains the pupils as per the methods outlined in self-management only and interdependent group contingency only, above. At the start of each lesson, a quick reminder of the rules is provided. Every 15 minutes, the teacher chooses at random, one of the classroom rules. Each pupil rates themselves on that rule, on a 0-4 scale. At the end of the lesson, the pupils then rate the whole class on a randomly chosen rule, on the same scale. The majority vote is taken as the class vote. Where the teacher's rating and the pupils' rating match, the pupils are awarded those points, plus a bonus point for matching the teacher's score (e.g. where the pupils and teacher both rated 4, the class would be awarded 5 points). Where there is a difference of one point between the pupils and teacher, the class would be awarded the teacher's points (e.g. where the teacher votes 2 and the pupil vote 3, the pupils are awarded 2 points). If the difference is more than one, no points are awarded. The points are marked on a graph in front of the pupils and will lead to the reward which was negotiated at the start of this intervention phase.

(Rule Reminder) Waitlist Control	The teacher trains the children on rule following and rule breaking behaviour only. The teacher begins the lesson every day by outlining/reminding the pupils of the classroom rules.
(No Change) Waitlist Control	The teacher conducts the lessons as usual. No training on the rules given.

Table 8.5: A table to outline the detailed procedures for the approaches used in this research.

8.11. Appendix: Teacher training slides for Phase A of the study



To refresh…

Class 1	Class 2	Class 3	Class 4
Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires
Self-management only	**Interdependent group contingency only**	**Rule reminder at the start of the lesson**	**Nothing different**
Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires
Interdependent group contingency + self-management	**Interdependent group contingency + self-management**	**Rule reminder at the start of the lesson**	**Nothing different**
Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires

To begin with…

- All teachers must agree 3 or 4 classroom rules for the students to follow, in each of the 3 classes.
- My suggestion: One about being positive and respectful of other pupils (to avoid any ganging up).
- 1)
- 2)
- 3)
- 4)

Understanding the rules

- **When introducing the rules, just before starting the first approach…**
- To ensure that the children understand what is rule breaking and rule following behaviour:
  - Class discuss examples of rule breaking and rule following behaviour.
  - Role play to illustrate rule breaking and rule following behaviour.
- **Display the rules visibly in the classroom**
- Waitlist control class – this should take around 15 minutes. The other classes – this is part of wider training given to the children.

Self-management

- Self-management aims to develop children's skills in regulating their own behaviour.
- **You will need:**

Rules on display in the classroom

Rating sheet

	Name: Date:	A	B	C	D
1					
2					
3					
4					
Whole class rating					
Rule:		0	1	2	3	4
		Not at all	Once or twice	Some times	Most of the time	All of the time

15 minutes

Self-management

First, training the children:

1) Introduce classroom rules (as discussed on previous slides)
2) Explain that the children will be stopped at specific times during the lesson. When stopped, the teacher will choose one rule at random and the children will be asked to rate their own behaviour since the last time they were stopped.
3) The ratings are: the behaviour was done 0= not at all; 1= some of the time; 2= half of the time; 3= most of the time; 4= all of the time.
4) Discuss and practise instances when a child would give themselves each of those ratings. Show them how to fill in the rating sheet.
5) **Most important:** encourage children to be honest in their ratings. Stress the importance for them to recognise where their true behaviour rating is. Tell them that they will not be in trouble for their ratings.
6) Explain that at the end of the lesson, the children will share their ratings with the person sitting next to them, to (1) talk about why they gave themselves that rating, and (2) to give feedback to their friend, to see whether they agree.
7) Explain that at the end of each lesson, they will rate the whole class on a randomly chosen rule using the 0-4 scale and this will be compared with the teacher's rating.

This training could take around 30 minutes or maybe the entire first lesson

Self-management procedures

1) Start the lesson by **reminding** the pupils of the **class rules.**
2) Ensure **each child has a rating sheet** with their name and date written on top.
3) Every **15 minutes** stop the class and choose a rule at random for them to rate themselves on **since the last time they were stopped/the start of the lesson**.
4) At the **end of the lesson**, choose a rule at random for the pupils to **rate the whole class** on.
5) The pupils rate the whole class on their sheet. The teacher rates the whole class on a separate sheet
6) The pupils put their hands up to **vote for their rating**. The majority is the class vote.
7) The **teacher reveals his/her rating** and gives feedback
8) Tell the pupils to **share their self-ratings** with their partner and justify their choice. Tell them to **give each other feedback** on whether they agree.

Interdependent group contingency

- Interdependent group contingency aims to improve behaviour through rewards and by increasing co-operation between pupils.
- **You will need:**

Interdependent group contingency

First, **training the children**:

1) Introduce classroom rules (as discussed on previous slides)
2) Explain to the children that the teacher will be keeping track of how well they as a whole class are following the rules. At the end of each lesson, the teacher will choose a rule at random and give the class a score between 0 and 4. When the class reach 50, they will receive a prize or reward.
3) Highlight the rule about the **importance** of being supportive, respectful and encouraging of each other, even if someone breaks the rule. Practise/roleplay with the children of how to do this.
4) **Agree with the children which reward(s) they would like** to collect points for. **Ensure that it is rewarding** for all pupils. An idea would be to collect a range of reward ideas and upon getting 50 points, one reward is picked at random.

This training could take around 20-30 minutes

Interdependent group contingency procedures

1) Start the lesson by **reminding** the pupils of the **class rules.**
2) Remind the pupils of the **rule to be supportive** and respectful of each other.
3) Remind children that they **need 50 points** for their class reward.
4) Teacher **keeps track of rule following and breaking** behaviour of the whole class.
5) At the end of the lesson, **randomly choose a rule**.
6) The **teacher awards the class** between 0-4 points and **marks this on the graph**.
7) The teacher **explains why the class received that rating** with clear examples of behaviour observed (without naming names).

Keeping a log

- To help me understand the data, please note down anything that you think might have an impact on the approaches. For example:
- Teacher absence so the approach not administered on a given day
- School trips
- Christmas play
- Days where parts of the procedure were missed out or went wrong.
- **Please include:**
  - What happened
  - What date it happened

Any questions?

Next steps:

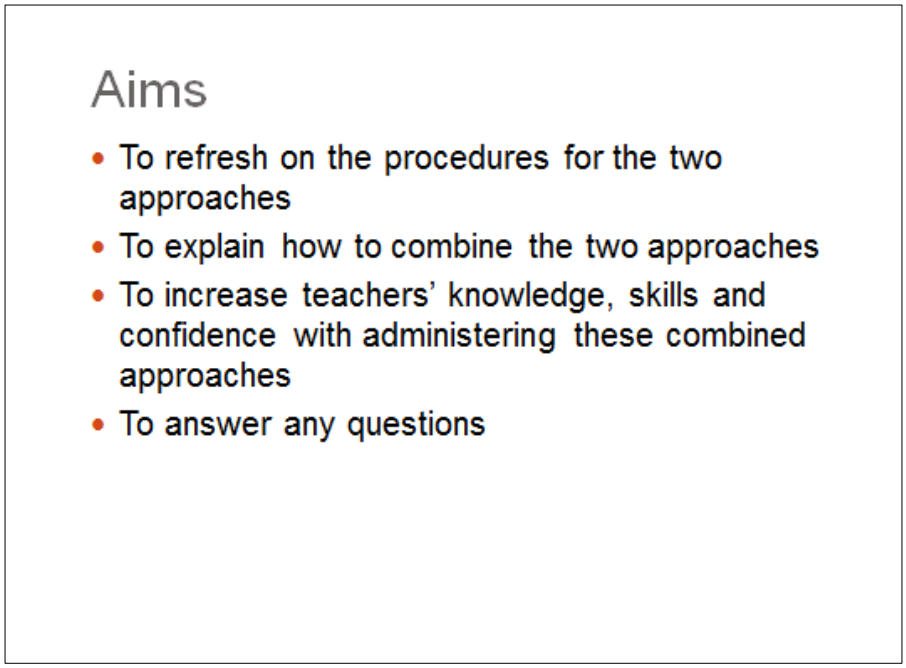
- Prep the resources
- Train your class on the rules
- Train your class on the procedures
- Run the approach for the next four school weeks
- Collect the pupil rating sheets (self-management), and teacher ratings (interdependent group contingency and self-management) in chronological order
- Keep a log of any important factors that might have affected the approaches
- I / Assistant Psychologists will conduct fidelity checks during those four weeks and support you
- Contact me if you have any questions or need help with anything
- I will return after four school weeks to do more observations, to give you the next set of questionnaires and to train you on the next stage; combining the approaches.

My contact details

- **At any point** if you need to contact me, please do:

- lpxknb@nottingham.ac.uk
- kbhana@**e.gov.uk
- **Office number:** **

8.12. Appendix: Teacher training slides for Phase B of the study



To refresh…

Class 1	Class 2	Class 3	Class 4
Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires
Self-management only	**Interdependent group contingency only**	**Rule reminder at the start of the lesson**	**Nothing different**
Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires
Interdependent group contingency + self-management	**Interdependent group contingency + self-management**	**Rule reminder at the start of the lesson**	**Nothing different**
Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires	Class observations & teachers fill out questionnaires

Combining the approaches

Training the children:

- Children who have received self-management, will require specific training in the interdependent group contingency aspects.
- Children who have received interdependent group contingency will require specific training in the self-management aspects.
- By the end, both classes will be trained on both approaches and the way the combined approach is run will be the **same in both classes.**

Children who need self-management training…

First, **training the children**:

1) Remind the children of the classroom rules
2) Explain that the children will be stopped at specific times during the lesson. When stopped, the teacher will choose one rule at random and the children will be asked to rate their own behaviour since the last time they were stopped.
3) The ratings are: the behaviour was done 0= not at all; 1= some of the time; 2= half of the time; 3= most of the time; 4= all of the time.
4) Discuss and practise instances when a child would give themselves each of those ratings. Show them how to fill in the rating sheet.

Continued…

1) **Most important:** encourage children to be honest in their ratings. Stress the importance for them to recognise where their true behaviour rating is. Tell them that they will not be in trouble for their ratings.
2) Explain that at the end of the lesson, the children will share their ratings with the person sitting next to them, to (1) talk about why they gave themselves that rating, and (2) to give feedback to their friend, to see whether they agree.
3) Explain that at the end of each lesson, they will rate the whole class on a randomly chosen rule using the 0-4 scale and this will be compared with the teacher's rating.
4) Explain that they will still earn points towards their group reward, but this time based on how well they match their ratings with the teacher's ratings (see 'How many points?' slide).

This training could take around 30 minutes or maybe the entire first lesson

Children who need interdependent group contingency training…

First, **training the children**:

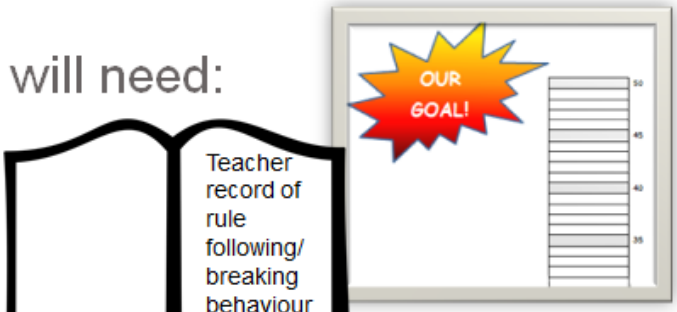
1) Remind the children of the classroom rules
2) Explain to the children that they will continue to rate their own and the class' behaviour on the sheet, and the teacher will also still rate the whole class as well.
3) However, explain that now, the class can earn points if their whole-class rating matches the teacher's rating of the class (see 'how many points?' slide and explain this to the children)
4) Explain that when the class reach 50, they will receive a prize or reward. **Agree with the children which reward(s) they would like** to collect points for. **Ensure that it is rewarding** for all pupils. An idea would be to collect a range of reward ideas and upon getting 50 points, one reward is picked at random.

Continued…

1) Highlight the rule about the **importance** of being supportive, respectful and encouraging of each other, even if someone breaks the rule. Practise/roleplay with the children of how to do this.

This training could take around 20-30 minutes

You will need:



Teacher record of rule following/ breaking behaviour

Rules on display in the classroom

Rating sheet

	Name: Date:	A	B	C	D
1					
2					
3					
4					
Whole class rating					
Rule:		0 1 2 3 4			

Combined approach procedures

- Teacher starts the lesson by reminding the pupils of the class rules
- Teacher reminds the pupils of the rule to be supportive and respectful of each other
- The contingency for group rewards are outlined: reaching 50 points means getting the reward.
- All pupils are provided with a rating sheet
- Every 15 minutes, the teacher chooses a rule for the pupils to rate themselves on their behaviour during the preceding 15 minutes
- The children reflect and rate themselves on the sheet
- At the end of the lesson, pupils share their self-ratings with their partner and justify their choice. Pupils give feedback to each other

Continued…

- At the end of the lesson, the teacher randomly chooses a rule for the pupils to rate the whole class on
- The pupils rate the whole class on the sheet. The teacher rates the whole class on a separate sheet.
- The pupils put their hands up to vote for their rating.
- The teacher shows his/her rating
- The teacher explains why the class received that rating with clear examples of behaviour observed (naming no names)
- The ratings are compared and the teacher awards the correct number of points to the class (**see next slide**)
- The teacher marks the points on a visual graph

How many points?

- Where the teacher's rating and the pupils' rating match, the children are awarded those points, plus a bonus point for matching the teacher's score (e.g. where the pupils and teacher both rated 4, the class would be awarded 5 points; 4 for the rating, plus an extra point for accuracy). Where there is a difference of one point between the pupils and teacher, the class would be awarded points based on the teacher's rating (e.g. where the teacher votes 2 and the children vote 3, the children are awarded 2 points). If the difference is more than one, no points are awarded. The points are marked on a graph which is clearly displayed in the classroom.

Important things to remember

- Continue to keep a log of anything that affects the intervention
- Ensure rating sheets have pupils' names and dates on
- Ensure the rating sheets and teachers' ratings are collected every day and filed in chronological order

Keeping a log

- To help me understand the data, please note down anything that you think might have an impact on the approaches. For example:
- Teacher absence so the approach not administered on a given day
- School trips
- English lessons being cancelled
- Days where parts of the procedure were missed out or went wrong.
- **Please include:**
  - What happened
  - What date it happened

Any questions?

Next steps:

- Prep the resources
- Train your class on the new procedures
- Run the combined approach for the next four school weeks
- Collect the pupil rating sheets and teacher ratings in chronological order
- Keep a log of any important factors that might have affected the approaches
- I will conduct fidelity checks during those four weeks and support you
- Contact me if you have any questions or need help with anything
- I will return after four school weeks to do the final observations and to give you the last set of questionnaires to fill out.

My contact details

- **At any point** if you need to contact me, please do:

- lpxknb@nottingham.ac.uk
- kbhana@**.gov.uk
- **Office number:** **

8.13. **Appendix: Resources used for the approaches**

Rating sheet

	Name: **Date:**	**A**	**B**	**C**	**D**
1	Be respectful to the teacher and to other children				
2	Stop and look at the teacher when the teacher is talking				
3	Follow teacher instructions straight away				
4	Allow other children to learn				

Whole class rating						
Rule:		0	1	2	3	4
		Not at all	Once or twice	Some times	Most of the time	All of the time

Our Class

Rule 1:

Be respectful to the teacher and to other children

Rule 2:

Stop and look at the teacher when the teacher is talking

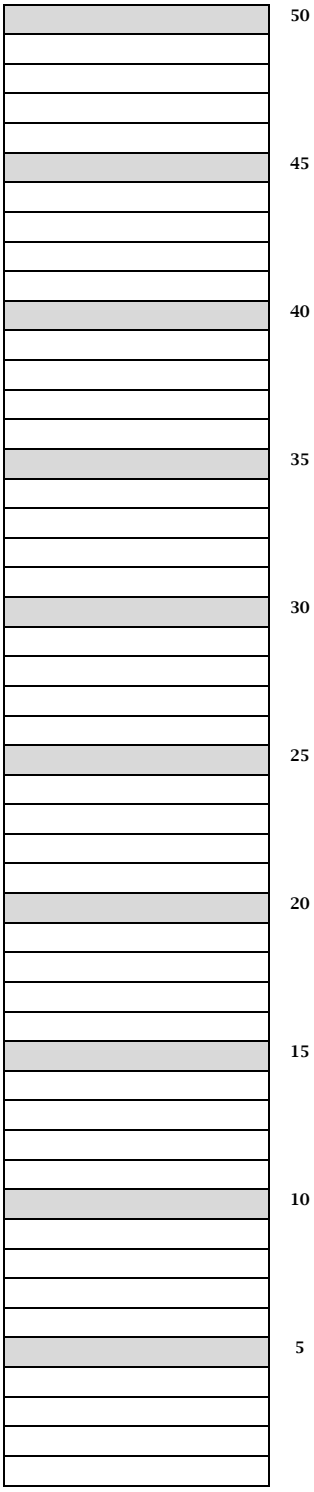
Rule 3:

Follow teacher instructions straight away

Rule 4:

Allow other children to learn


OUR
GOAL!


50
45
40
35
30
25
20
15
10
5

8.14. **Appendix: Teacher record sheet**

Teacher record of class ratings

0= not at all; 1= some of the time; 2= half of the time; 3= most of the time; 4= all of the time

Date	Lesson/subject	Rule (number)	Teacher's whole class rating (number)

8.15. **Appendix: Fidelity checklists**

Self-management

1	Teacher starts the lesson by reminding the pupils of the class rules	YES/NO/PARTIALLY
2	All pupils are provided with a rating sheet	YES/NO/PARTIALLY
3	Every 15 minutes, the teacher chooses a rule for the pupils to rate themselves on their behaviour during the preceding 15 minutes	YES/NO/PARTIALLY
4	The children reflect and rate themselves on the sheet	YES/NO/PARTIALLY
5	At the end of the lesson, the teacher randomly chooses a rule for the pupils to rate the whole class on	YES/NO/PARTIALLY
6	The pupils rate the whole class on the sheet	YES/NO/PARTIALLY
7	The teacher rates the whole class on a separate sheet	YES/NO/PARTIALLY
8	The pupils put their hands up to vote for their rating	YES/NO/PARTIALLY
9	The teacher shows his/her rating	YES/NO/PARTIALLY
10	Pupils share their self-ratings with their partner and justify their choice	YES/NO/PARTIALLY
11	Pupils give feedback to each other	YES/NO/PARTIALLY
Comments/notes		

Interdependent group contingency

1	Teacher starts the lesson by reminding the pupils of the class rules	YES/NO/PARTIALLY
2	Teacher reminds the pupils of the rule to be supportive and respectful of each other	YES/NO/PARTIALLY
3	The contingency for group rewards are outlined: reaching 50 points means getting the reward.	YES/NO/PARTIALLY
4	Teacher keeps track of rule following and breaking behaviour of the whole class	YES/NO/PARTIALLY
5	At the end of the lesson, the teacher chooses a rule at random	YES/NO/PARTIALLY
6	The teacher awards the class between 0-4 points	YES/NO/PARTIALLY
7	The teacher explains why the class received that rating with clear examples of behaviour observed	YES/NO/PARTIALLY
8	The teacher marks the points on a visual graph	YES/NO/PARTIALLY
Comments/notes		

Fidelity Check

Self-management and interdependent group contingency

1	Teacher starts the lesson by reminding the pupils of the class rules	YES/NO/PARTIALLY
2	Teacher reminds the pupils of the rule to be supportive and respectful of each other	YES/NO/PARTIALLY
3	The contingency for group rewards are outlined: reaching 50 points means getting the reward.	YES/NO/PARTIALLY
4	All pupils are provided with a rating sheet	YES/NO/PARTIALLY
5	Every 15 minutes, the teacher chooses a rule for the pupils to rate themselves on their behaviour during the preceding 15 minutes	YES/NO/PARTIALLY
6	The children reflect and rate themselves on the sheet	YES/NO/PARTIALLY
7	At the end of the lesson, pupils share their self-ratings with their partner and justify their choice	YES/NO/PARTIALLY
8	Pupils give feedback to each other	YES/NO/PARTIALLY
9	At the end of the lesson, the teacher randomly chooses a rule for the pupils to rate the whole class on	YES/NO/PARTIALLY
10	The pupils rate the whole class on the sheet	YES/NO/PARTIALLY
11	The teacher rates the whole class on a separate sheet	YES/NO/PARTIALLY
12	The pupils put their hands up to vote for their rating	YES/NO/PARTIALLY
13	The teacher shows his/her rating	YES/NO/PARTIALLY
14	The teacher explains why the class received that rating with clear examples of behaviour observed	YES/NO/PARTIALLY
15	The ratings are compared and the teacher awards the correct number of points to the class	YES/NO/PARTIALLY
16	The teacher marks the points on a visual graph	YES/NO/PARTIALLY

Comments/notes

8.16. **Appendix: Outcomes of the fidelity checks**

Class observed	Date	Observer	Activity	Outcome
Self-management	01.11.16	Researcher & Assistant 1	Training class on the approach	Done to fidelity
	03.11.16		Running the approach	Mostly done to fidelity. Feedback for improvement given.
	15.12.16	Researcher	Training class on the approach	Done to fidelity
	12.01.16	Researcher	Running the approach	Mostly done to fidelity. Rule reminder forgotten and peer feedback. Feedback for improvement given.
	19.01.17	Researcher	Running the approach	Mostly done to fidelity. Missed one rating opportunity and forgot peer feedback. Feedback for improvement given
Interdependent group contingency	01.11.16	Assistant 2	Training class on approach	Done to fidelity
	03.11.16		Running the approach	Done to fidelity

	27.01.17	Researcher	Training class on the approach. Running the approach	Done to fidelity. Just reminded to explain why teacher rated the class a particular score.
	23.02.17	Researcher	Running the approach	Done to fidelity
No change (waitlist control) class	25.11.16	Assistant 1	Teaching as usual	No diffusion of treatment.
	11.01.16	Assistant 1	Teaching as usual	No diffusion of treatment
Rule reminder (waitlist control) class	01.11.16	Assistant 3	Training class on approach	Done to fidelity
	03.11.16		Running the approach	Done to fidelity. No diffusion of treatment.
	11.01.16	Assistant 3	Running the approach	Done to fidelity. No diffusion of treatment.

Table 8.6: A table to summarise the outcomes from the fidelity checks conducted in this research.

8.17. Appendix: Strengths and Difficulties Questionnaire

<Removed due to copyright restrictions>

<Removed due to copyright restrictions>

8.18. Appendix: Observation schedule

Individuals – fixed order – 15 sec/5 sec

Class
Lesson
Date

Time
T1 / T2 / T3
Percentage off-task

On task?	Agree?	On task?	Agree?	On task?	Agree?	On task?	Agree?
TOTAL	/		/		/		/

8.19. **Appendix: Debrief letter**

School of Psychology
Debrief Sheet

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: **The Impact of Self-Management and Group Contingency Approaches on Whole-class Behaviour**
Ethics Approval Number or Taught Project Archive Number: ***820***
Researcher: ***Kamal Bhana email: lpxknb@nottingham.ac.uk***
Supervisor: ***Dr Nathan Lambert email: nathan.lambert@nottingham.ac.uk***

Dear (Parent/carer/teacher/head teacher/governor),

Thank you for participating in my research project. This study aimed to evaluate how effective two whole-class approaches for classroom management were in increasing engagement in the classroom.

In order to research this, classes either received self-management and group contingency or no approaches. This allowed me to compare the impact of these approaches. I asked teachers to complete questionnaires on all of the children except those who were opted-out and I also observed the children in class a few times.

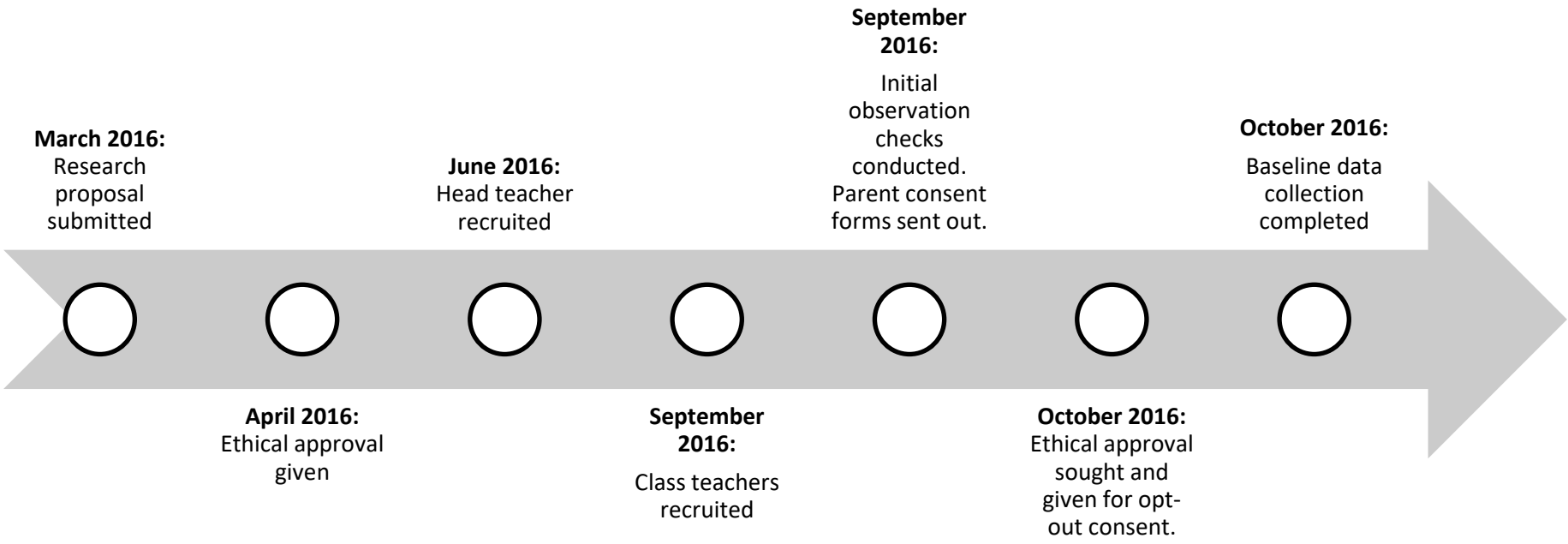
The findings from this research suggest that the classes who used these approaches showed less overall off-task behaviour in the classroom, than at the start of the study. The classes that did not receive the approaches during the study will be offered the training.

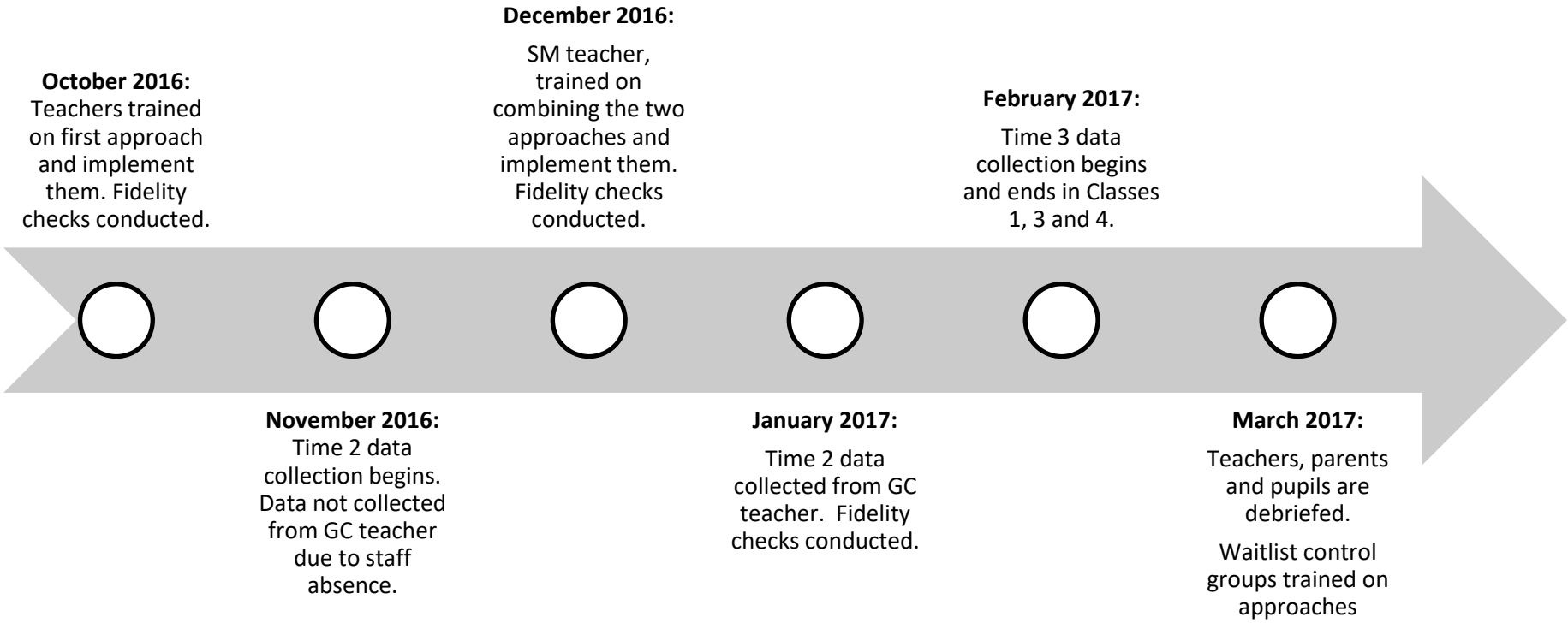
If you have any questions about the research, I will be available to meet with you at the school on **Friday 24th March 2017,** at **2.30pm** or **4.15pm**. Alternatively, please contact me via email.

Yours sincerely,

Kamal Bhana
Trainee Educational Psychologist

8.20. Appendix: Timeline of Research Process




October 2016:
Teachers trained on first approach and implement them. Fidelity checks conducted.
November 2016:
Time 2 data collection begins. Data not collected from GC teacher due to staff absence.
December 2016:
SM teacher, trained on combining the two approaches and implement them. Fidelity checks conducted.
January 2017:
Time 2 data collected from GC teacher. Fidelity checks conducted.
February 2017:
Time 3 data collection begins and ends in Classes 1, 3 and 4.
March 2017:
Teachers, parents and pupils are debriefed.
Waitlist control groups trained on approaches

8.21. **Appendix: Alpha level correction details**

Holm's (1979) procedure for adjusting the alpha (p level) and checking significance:

- Order the reported p statistics from smallest to largest.
- Start with the smallest p statistic and calculate α/T (Holm's adjusted probability level) where α is .05 (the p level generally used in social science research), and T refers to the number of tests conducted (which in this research is 24). If Holm's adjusted p level is larger than the reported p level, there is significance.
- Move to the next p statistic and calculate $\alpha/(\mathrm{T}-1)$ and check again for significance.
- For the next statistic, calculate $\alpha/(\mathrm{T}-2)$ and so on, down the list of reported p levels.

The table highlights that using Holm's correction and a more conservative alpha of p=.01 both identify the same test results as showing statistical significance.

Research Question	**Statistical Test used**	**Comparisons**	**Reported *p* level**	**Significant when using *p*=.01?**	***P* level correction using Holm's procedure**	**Significant when using Holm's correction?**
1/2	Kruskall-Wallis	T2 – hyperactivity and concentration – compare all classes	0	Yes	0.00217	Yes
1/2	Mann-Whitney U	T2 – hyperactivity and concentration – Class 1 v Class 4	0	Yes	0.00227	Yes
1/2	Mann-Whitney U	T2 – hyperactivity and concentration – Class 2 v Class 4	0	Yes	0.00238	Yes

Research Question	Statistical Test used	Comparisons	Reported *p* level	Significant when using *p*=.01?	*P* level correction using Holm's procedure	Significant when using Holm's correction?
3	Mann-Whitney U	T2 – hyperactivity and concentration – Class 1 v Class 2	0	Yes	0.0025	Yes
4	Kruskall-Wallis	T3 – hyperactivity and concentration – compare classes 1, 3, 4	0	Yes	0.00263	Yes
1/2	Mann-Whitney U	T2 – hyperactivity and concentration – Class 2 v Class 3	0.001	Yes	0.00277	Yes
3	Wilcoxon Signed Rank	Class 2 – behaviour – T1 –-> T2	0.001	Yes	0.00294	Yes
3	Wilcoxon Signed Rank	Class 1 – hyperactivity and concentration – T1 –-> T2	0.001	Yes	0.00312	Yes
3	Wilcoxon Signed Rank	Class 2 – hyperactivity and concentration – T1 –-> T2	0.001	Yes	0.00333	Yes
4	Kruskall-Wallis	T2 – hyperactivity and concentration –compare classes 1, 3, 4	0.002	Yes	0.00357	Yes
1/2	Kruskall-Wallis	T2 – behaviour – compare all classes	0.011	No	0.00384	No
4	Kruskall-Wallis	T3 – behaviour – compare classes 1, 3, 4	0.033	No	0.00416	No
1/2	Kruskall-Wallis	T1 – hyperactivity and concentration –compare all classes	0.057	No	0.00454	No

Research Question	Statistical Test used	Comparisons	Reported *p* level	Significant when using *p*=.01?	*P* level correction using Holm's procedure	Significant when using Holm's correction?
4	Wilcoxon Signed Rank	Class 3 – behaviour – T2 –-> T3	0.119	No	0.005	No
4	Wilcoxon Signed Rank	Class 3 – hyperactivity and concentration – T2 –-> T3	0.169	No	0.00555	No
4	Kruskall-Wallis	T2 – behaviour – compare classes 1, 3, 4	0.25	No	0.00625	No
1/2	Mann-Whitney U	T2 – hyperactivity and concentration – Class 1 v Class 3	0.469	No	0.00714	No
4	Wilcoxon Signed Rank	Class 4 – hyperactivity and concentration – T2 –-> T3	0.496	No	0.00833	No
1/2	Kruskall-Wallis	T1 – behaviour – compare all classes	0.544	No	0.01	No
4	Wilcoxon Signed Rank	Class 4 – behaviour – T2 –-> T3	0.596	No	0.0125	No
3	Wilcoxon Signed Rank	Class 1 – behaviour – T1 –-> T2	0.719	No	0.01666	No
4	Wilcoxon Signed Rank	Class 1 – behaviour – T2 –-> T3	0.813	No	0.025	No
4	Wilcoxon Signed Rank	Class 1 – hyper – T2 –-> T3	0.833	No	0.05	No

Table 8.7: A table to show which statistical tests yielded statistically significant results according to Holm's (1979) correction and as a result of employing a p=.01 level of significance.

8.22. **Appendix: Inter-rater reliability check sheet**

Inter-rater reliability for Phase A

Scale
1 = Definite deterioration
2 = Slight deterioration
3 = Definite no change
4 = Slight improvement
5 = Definite improvement
6 = Unsure

	Researcher	**Colleague**
CLASS 1		
• *What change has been observed between Time 1 and Time 2 for off-task behaviour?*		
• *What change has been observed between Time 1 and Time 2 for disruptive behaviour?*		
CLASS 2		
• *What change has been observed between Time 1 and Time 2 for off-task behaviour?*		
• *What change has been observed between Time 1 and Time 2 for disruptive behaviour?*		
CLASS 3		
• *What change has been observed between Time 1 and Time 2 for off-task behaviour?*		
• *What change has been observed between Time 1 and Time 2 for disruptive behaviour?*		
CLASS 4		
• *What change has been observed between Time 1 and Time 2 for off-task behaviour?*		
• *What change has been observed between Time 1 and Time 2 for disruptive behaviour?*		

Inter-rater reliability for Phase B

Scale

1 = Definite deterioration
2 = Slight deterioration
3 = No change
4 = Slight improvement
5 = Definite improvement
6 = Unsure

	Researcher	**Colleague**
Class 1		
• *What change has been observed between Time 2 and Time 3 for off-task behaviour?*		
• *What change has been observed between Time 2 and Time 3 for disruptive behaviour?*		
Class 3		
• *What change has been observed between Time 2 and Time 3 for off-task behaviour?*		
• *What change has been observed between Time 2 and Time 3 for disruptive behaviour?*		
Class 4		
• *What change has been observed between Time 2 and Time 3 for off-task behaviour?*		
• *What change has been observed between Time 2 and Time 3 for disruptive behaviour?*		