UNIVERSITY OF NOTTINGHAM

The impact of training in a pupil centred behaviour plan on staff self-efficacy, staff burnout, and pupil challenging behaviour.

by Heather Cooke

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

July 2014

Acknowledgements

The author would like to thank:

The parents, students, and school staff for taking part in, and dedicating their time and effort to this research.

Matthew for his dedication to the young people he works with, resulting in his proactive development of the behaviour plan. Also for his time and support in recruiting participants and developing training.

The Local Authority, in particular Sharon and Gillian, for making the research possible and providing me with the time and support required to complete this project.

The Nottingham University tutor team for their support, teaching, and guidance throughout the three years of the doctoral course.

Family, friends, and Team TEP for their constant encouragement, support, and patience.

Abstract

Challenging behaviour in schools is a phenomenon focused on by a number of educational documents (Ofsted, 2010) and the media (Vasager, 2011). Challenging behaviour has been shown to have negative impact on a number of student and staff outcomes (DfE, 2012a). Staff outcomes impacted by challenging behaviour include increasing burnout (Crone, Hawken & Bergstrom, 2007) and decreasing self-efficacy (Mitchell & Hastings, 2001), which have been connected to negative impact on staff health (Hastings & Bham, 2003). Time allocated to staff training in schools is decreasing (Bubb & Earley, 2013), highlighting a need for research which considers how school staff can be supported in the limited time available.

The impact of a behaviour plan based on solution focused and behavioural principles (developed by a specialist teacher) on students' challenging behaviour was explored through single case experimental design. Further to this a randomised control design investigated the impact of whole school training relating to the behaviour plan on school staff burnout and self-efficacy. Results showed that the intervention reduced challenging behaviours to differing degrees of all 3 primary students included in the single case experimental design. The whole school training did not significantly impact the number of behaviour plans implemented in classrooms. However, training was shown to have significant positive effects on school staffs' personal, general, and overall self-efficacy, with no impact on external self-efficacy. The training was also shown to significantly decrease school staffs' burnout levels; specifically physical fatigue, cognitive weariness, and overall burnout levels, but not emotional exhaustion.

This research suggests that the behaviour plan and the accompanying whole school training have the potential to increase teacher self-efficacy, decrease teacher burnout, and provide staff with a suitable intervention to manage challenging behaviour. Areas for further research are highlighted by the limitations and additional observations made during the research process.

Table of Contents

Appendices Contents

Table of Figures

List of Tables

List of Acronyms and Abbreviations

Abbreviation/ Acronym	Definition
ABA	Applied Behavioural Analysis
ATA	Advanced Teaching Assistant
BEMS	Behaviour Enhanced Mainstream Service
DAppEdPsy	Doctorate of Applied Educational Psychology
Df	Degrees of Freedom
DFE	Department for Education
DRA	Differential Reinforcement of Alternative Behaviours
DRI	Differential Reinforcement of Incompatible Behaviours
DRL	Differential Reinforcement of Lower Behaviour Rates
DRO	Differential Reinforcement of Other (Omission of) Behaviours
EP	Educational Psychologist
EPS	Educational Psychology Service
FBA	Functional Behavioural Analysis
GTA	General Teaching Assistant
INSET	In Service Education and Training
MBI	Maslach Burnout Inventory
NQT	Newly Qualified Teacher
OFSTED	Office for Standards in Education
PCP	Person Centred Psychology
RCT	Randomised Control Trial
SCED	Single Case Experimental Design
sd	Standard Deviation

SEN	Special Educational Needs
SENCO	Special Educational Needs Coordinator
SFP	Solution Focused Psychology
SFBT	Solution Focused Brief Therapy
SFT	Solution Focused Therapy
SMBM	Shirom-Melamed Burnout Measure
TA	Teaching Assistant
TEACCH	Treatment and Education of Autistic and Related Communication Handicapped Children
TECDMS	Teacher Efficacy in Classroom Management and Discipline Scales
TEP	Trainee Educational Psychologist
UK	United Kingdom

1 Literature Review

1.1 Introduction

This introduction will present a brief summary of key issues which are further developed within the following literature review and research. The researcher has a personal and professional interest in the area of challenging behaviour, having worked in a variety of educational settings where negative impacts of challenging behaviour were witnessed both on student educational outcomes, and staff well-being. The settings the researcher worked in highlighted that the different levels of training staff receive is dependent on setting idiosyncrasies. This enhanced the researcher's interest in the potential benefits of staff training for school staff and students. The researcher's interest in the evidence base of staff training and potential outcomes has been further cemented through her current role as a trainee educational psychologist (TEP), which includes development and delivering of training for school staff.

Challenging behaviour has been a long-standing issue raised regularly in educational documents (Ofsted, 2010) and the media (Vasager, 2011). Over recent years the media has reported an increase of challenging behaviour in schools (Sellgren, 2013), and statistics showed an increase in the number of permanent exclusions from 2010-2011 to 2011-2012 ((Department for Education (DfE) 2013)), with permanent exclusions in primary schools increasing by 13.9% in this time period (58% of these primary school exclusions were due to physical or verbal abuse). Domokos (2012) also warned that these national statistics do not fully encompass the difficulties in schools as they do not include the high number of unofficial exclusions and managed moves. Hayden (2014) conducted research over the past ten years with staff and schools and reports that the Government documents reporting an improvement in challenging behaviour since 2012 (with Ofsted reports showing 97% of schools have satisfactory or good behaviour (DfE, 2012a)) does not reflect the actual situation within classrooms where behaviour difficulties are still a consistent difficulty for teachers. It appears that challenging behaviour poses continuing difficulties for teachers, which, when not managed effectively often have extensive negative impacts on those students exhibiting challenging behaviour, their peers, and staff health (DfE, 2012a).

Challenging behaviour can create difficulties in and out of school for both students and educational staff. Challenging behaviour has been shown to impact negatively on school staff coping in relation to burnout (Mitchell & Hastings, 2001) and self-esteem (Crone et al, 2007), ultimately impacting on staff health (Hastings & Bham, 2003), leading to staff leaving the profession (Kokkinos, 2007). This can then lead to students experiencing less staff consistency, potentially increasing negative outcomes for students as the profession continues to lose trained professionals. This suggests that despite long-standing research into behaviour management, there are a number of difficulties relating to challenging behaviour that the education system is still not managing to overcome, thus creating a negative cycle for students and professionals. This cycle suggests a continued need for furthering staff knowledge, resiliency, and ability to manage student challenging behaviour. Consequently, it is not only important that training provided for managing challenging behaviour leads to implementation of strategies that effectively decrease student challenging behaviour, but also has a direct positive impact on staff coping.

Knowledge about challenging behaviour has been shown to increase teacher self-efficacy (Brinson & Steiner, 2007) and reduce teacher burnout (Mitchell & Hastings, 2001). Self-efficacy and burnout have been found to be significant predictors of staff health (Hastings & Bham, 2003) and attrition (Weisberg & Sagie, 1999). This research considers how staff outcomes of training may be as crucial as student outcomes, as reduced negative affect of teachers towards students displaying challenging behaviour has been shown to increase behaviour management skills and student academic outcomes (Mitchell & Hastings, 2001).

With consideration to the impact challenging behaviour is reported to have on all students (Ofsted, 2010), challenging behaviour presents itself as a key area for further research in the educational field. Information and training in behaviour management during teacher training is limited, and this was highlighted as an area of need by the DfE in 2012 (DfE, 2012c). This has led to the development of documents such as "Improving Teacher Training for Behaviour" from the Teaching Agency that highlighted

good practice in teacher training courses where behaviour is a focus (Teaching Agency, 2012). The current increase of quicker teacher training routes in the UK (e.g. Teach First trainees are based full-time in a classroom after 6 weeks (Teach First, 2014, online)), is likely to reduce the time for training regarding behaviour management further. Schools and "mentors" are then required to support staff, relying on In-Service Education and Training (INSET) days and after-school sessions for further training ((Teach First, 2014, online)). However, there is limited research on how effective these sessions are, as training literature suggests training should not be limited to "one stop trainings" (Barton, Chen, Pribble, Pomes & Kim, 2013, p344), but needs to include long-term support. Unfortunately due to the educational climate in the UK and time restraints in schools, this type of long-term teacher training is unlikely in many schools (Bubb & Earley, 2013). This research therefore aims to explore the effectiveness of the distribution of information and knowledge in after school training sessions, relating to managing challenging behaviour in class, highlighted as a continuing difficulty for staff (Hayden, 2014). Both staff self-efficacy and burnout ratings have been linked to student outcomes and teacher health (Hastings & Bham, 2003), emphasising the importance of research that considers how to impact these teacher constructs.

1.2 Challenging Behaviour in Schools

1.2.1 What is Challenging Behaviour?

The media portrays challenging behaviour as an increasing phenomenon in schools (Sellgren, 2013), as do a number of governmental documents (DfE, 2010). However, challenging behaviour has been defined in a variety of ways throughout literature (DfE, 2010). The term 'challenging behaviour' was promoted by The Association for People with Severe Handicaps in the 1990's with the aim of replacing terms such as abnormal, aberrant, disturbed, dysfunctional, and maladaptive (Emerson, 2001). It was felt that the term 'challenging behaviour' provided an alternative descriptor free from implicit assumptions about psychological characteristics of behaviour and the person engaging in this behaviour (Emerson, 2001). Emerson (2001) defined challenging behaviour as referring to behaviours which;

"involve significant risks to people's physical well-being or act to reduce markedly access to community settings" (pg 3).

Contrastingly, other definitions include specific descriptions of behaviours ranging from "extreme acts of violence" to "low-level frequent disruption that is the most common form of pupil misbehaviour" (DfE, 2012a). The DfE, 1994 describe emotional and behavioural disorders as ranging:

"…from social maladaptation to abnormal emotional stresses. They are persistent … and constitute learning difficulties. They may be multiple and may manifest themselves in many different forms and severities. They may become apparent through withdrawn, passive, aggressive or self-injurious tendencies." (DfE, 1994, pg7)

In a review of challenging behaviour Cameron (1998) reported five behaviour classifications: aggressive behaviour; physically disruptive behaviour; socially disruptive behaviour; authority-challenging behaviour; self-disruptive behaviour. There are often disagreements in the literature about which behaviours are the most disruptive

and cause the most difficulties for teachers (DfE, 2010). Watkins & Wagner (2000) report that teachers describe low-level disruptive behaviour as most difficult, due to its frequency.

Watkins & Wagner (2000) report that the definition of challenging behaviour depends on the teacher, setting, expectations, and history of actor and observer. This corresponds with the contextualist world view of behavioural analysis that defines challenging behaviour as a social construction (Emerson, 2001). Challenging behaviour has been described as a response to challenging situations, being defined by the social consequences of the behaviour in the setting, and according to the people involved (Emerson, 2001). The following sections will summarise and highlight key research in areas that are directly relevant to the research.

1.2.2 Interventions and Strategies to Manage Challenging Behaviour

A variety of methods of behaviour management stem from a number of psychological fields, with professional preferences for methods changing over time. In a review of behavioural research, Dunlap et al, (2006) concluded that in order to successfully alter challenging behaviour, interventions need to include: teaching replacement skills; information from functional analyses of behaviour; adjustment of antecedent conditions; a number of environmental settings; and input from family members. Strategies recommended in governmental guidance and used by schools in the UK include the use of seating plans, consistent rewards and sanctions, clear exclusion policies, recognition of pupil voice (Steer, 2009), rewarding good behaviour, clear consequences for negative behaviour, and on-site out of class provision (DfE, 2012a).

A review of behavioural interventions in primary schools in the UK (Evans, Harden, Thomas & Benefield 2003) discusses a number of strategies currently used to manage challenging behaviour. Evans et al, (2003) report there is a notable lack of literature evaluating the effectiveness of behavioural interventions used in schools. However, the most commonly used behaviour management strategies tend to stem from the psychological principles of behavioural, cognitive behavioural, and systemic psychology (Evans et al, 2003). The most commonly used and applied strategies tend to

be based on principles of behavioural psychology (Evans et al, 2003). Recent governmental documents have focused on managing challenging behaviours through increased teacher control and punishment (DfE, 2012b).

1.2.3 Behavioural Psychology (Functions)

Behavioural psychology is often used to support interventions and guide strategies for managing challenging behaviour. Behavioural psychology states that all behaviour is learned through contingent reinforcement and punishment (see 1.2.6 for definitions) in a person's environment (Skinner, 1971) and impacted by the systems surrounding the individual (Fox, 2009). Behaviour has been said to serve one or more of five behavioural functions: social attention; tangible reinforcement; self-stimulation; social negative reinforcement (escape); and automatic negative reinforcement (removal of aversive stimuli) (Cooper et al, 2007). It is recommended that considerations of antecedents and consequences of behaviours guide hypothesis formulation about behavioural functions (Cooper et al, 2007). Functional behavioural analysis (FBA) involves identifying the functions of behaviour through exploration of antecedents and consequences, strategies and interventions should then be based on these functions to address the behaviours (Cooper et al, 2007). The Royal College of Psychiatrists et al (2007) produced guidance dictating that all challenging behaviours should always be assessed through FBA in order to lead to treatment plans.

FBA is intended to inform strategies to manage challenging behaviour by identifying areas of need for replacement behaviours and any pre-requisite skills required for these replacement behaviours. FBA considers antecedents of behaviour that can be environmental (external), related to the student's own personal skills and preferences (internal), or an interaction between these external and internal influences, and considers how to manage these within relevant environments. FBA must be relevant to the student, and lead to effective application that is generalised across their natural environments (Cooper et al, 2007). To ensure generalisation it is necessary to consider both internal and external antecedents and consequences to the behaviour. Socially appropriate behaviours that aim to meaningfully fulfil functions originally being met through challenging behaviour are taught and reinforced based on the FBA (Cooper et

al, 2007). Behavioural psychologists believe that unless the challenging behaviour is decreased, the function of the original target behaviour cannot be being fulfilled, therefore the FBA is not accurate or complete (Gable & Hendrickson, 2000). There have been numerous studies showing behavioural interventions based on FBA successfully decrease challenging behaviours in mainstream settings (Goh & Bambara, 2010). However, research has found that the implementation of FBA in schools can be complex and time consuming, FBA is therefore often deemed inappropriate for mainstream settings (Moreno, 2010). Some research and psychological theories believe FBA simplifies behaviour change as it does not encourage self-management and awareness and does not incorporate understanding of within child abilities and attributions, consequently recommending a multi- dimensional approach (Gable & Hendrickson, 2000).

A recent systematic review of research, conducted between 1990–2007 evaluating school based interventions using FBA found that these interventions successfully decreased challenging behaviour in schools (Wood, Cho Blair & Ferro, 2009). McIntosh, Brown & Borgmeister (2008) conducted a meta-analysis of single case experimental design research and found that interventions based on FBA were more successful than those that did not include FBA. However, due to the heterogeneous nature of people engaging in challenging behaviour, very few studies use group methodologies. There is also a limited number of studies that have explored maintenance or generalisation of behaviour improvements stemming from FBA. However, Moreno (2010) concluded that interventions targeting challenging behaviour, informed by FBA, are more successful in their longitudinal outcomes than those not based on FBA.

It is recommended that in order to conduct FBA, discussions should focus on "what" and "when" questions addressing environmental factors relating to specific, measurable, and observable behaviours, as well as observational records of antecedents and consequences (Cooper et al, 2007). Mossmann, Hastings & Brown (2002) suggested that school staffs' negative emotional responses to young people with challenging behaviour decreases as their understanding of the functions of the challenging behaviour

increases. Consequently, in schools, staff speaking with pupils about the function of their behaviour, and attempting to explore these issues may potentially increase teacher ability to manage the behaviour with fewer emotions (Hastings & Brown, 2002). De Courcey-Bower, (2011) showed that increasing the number of links between FBA and behaviour plans results in greater behavioural change.

A review of research considering important factors for FBA concluded that in order to achieve effective and sustainable behaviour change, factors that maintain / reinforce the behaviour including systemic factors must be considered (Sugai et al, 2000). This is related to the view that behavioural functions are not necessarily fixed traits but are due to interactions between the environment and the individual (McIntosh et al, 2007).

1.2.4 Solution Focused Questioning and Challenging Behaviour

Systems theory can often inform interventions managing challenging behaviour through the use of solution focused approaches (Pelligrini, 2009). Solution focused therapy (SFT) aims to consider interactions between individuals, recognise the individual's skills and successful strategies they have applied in the past, and utilise these to effect change (Porter, 2007). Solution focused psychology (SFP) approaches assume that clients know their own problems best and are therefore capable of generating their own solutions (Miller & de Shazer, 2000). SFP is believed to impact clients as it encourages them to develop a new and positive story for themselves. This more positive story is developed by focusing on the client's own skills, strengths, and resources (Rae & Daly, 2008). There is a distinct lack of research evaluating the impact of SFP working within educational settings (Pelligrini, 2009). The majority of the research that does exist usually uses solution focused brief therapy (SFBT) in isolation.

Due to the small amount of empirical research evaluating the impact of SFBT (Corcoran, 2006) the evidence in this area is limited. A meta-analysis of SFBT suggests small but positive effects of SFBT on internalised behaviours in 12 studies in clinical and educational settings (Kim, 2008). However, no effect was found in the 9 studies that explored the impact of SFBT on externalised behaviours (Kim, 2008). Contrastingly, a later review that used only school based research (only 4 studies) found SFBT to decrease externalised behaviours with a medium to large effect size (Kim & Franklin, 2009). The review also found mixed impacts of SFBT on academic and attendance outcomes (Kim & Franklin, 2009). Through a randomised control trial (RCT) Daki & Savage (2010) found SFBT homework support had a significantly greater positive effect on participant general intelligence, reading, writing, and self-appraised behaviour conduct than academic homework support. The authors concluded this was due to the focus of SFBT on meta-cognitive skills allowing them to better access their academic work. While Corcoron (2006) found no difference in behaviour using SFBT compared to cognitive behavioural therapy (although both had positive effects), SFBT led to increased engagement resulting in a lower drop-out rate. It may therefore be that some of the positive effects reported in the literature about SFBT may be due to indirect effects on thought processes and ways in which it impacts one's ability to approach

tasks / situations. SFP incorporates a variety of techniques (de Shazer & Berg, 1997) such as looking for solutions and goal setting.

1.2.5 Student Led Behaviour Plans

As previously discussed (see 1.2.4), goal setting and homework are part of the components of SFT (de Shazer & Berg, 1997). These approaches concur with current government guidance, stating that students should always be involved with review meetings, and their voices considered and included in educational decisions (SEN Code of Practice, DfES, 2001). The focus on involving pupils in their own decision making has been increased in the recent Children and Families Bill, 2013.

The theory that students benefit from setting their own targets and taking control of their own behaviour plan is based on theories guiding solution focused consultations; a client knows his / her own skills best, so is therefore best able to generate solutions (Miller & de Shazer, 2000). Johnson & White (1971) summarised this effect suggesting that approaches where clients are relied upon to focus on their own strengths and targets allow clients to develop a positive self-image, leading to increased maintenance of behavioural changes in the absence of explicit external evaluations and feedback. Reese, Sherman, & Sheldon (1984) suggested that by shifting the "locus of control" to the client, their self-image becomes more positive, becoming the internal reinforcer, therefore increasing generalisation of skills as their positive self-image develops.

The idea of client involvement in their own treatment plans also stems from the humanistic psychological approach of person centred psychology (PCP) (Joseph & Murphy, 2012). PCP has at its core the basic assumption that people are "intrinsically motivated towards optimal positive psychological approaches" (Joseph & Murphy, 2012). This perspective suggests that humans have a biological imperative to grow proactively towards autonomy, but this relies on "optimal environmental conditions" (Joseph & Murphy, 2012). Beadle-Brown, Hutchinson & Whelton (2012) conducted research in group homes for adults with severe and profound learning difficulties. They used a method of support based on PCP; "active support". Active support aims to increase choice making opportunities to give clients increased control in their lives. It

was found that the use of active support increased the participants' engagement in their care as well as correlating with a reduction in challenging behaviours. This suggests that by involving clients in their own care plans and allowing them an increased number of choices, their engagement is increased, which in turn leads to increased success within their environments. Children have been shown to have the ability to consider a variety of problems and solutions strategically and effectively when guided through the process (Elliot & Faupel, 1997). Strategies such as the use of PCP with young people have increased in popularity with educational psychologists (EPs) (Lokke, Gersch, M'gadzah & Frederickson, 1997) as they are able to ascertain views of the young person, therefore increasing engagement (Beadle-Brown et al, 2012) and the likelihood of setting achievable targets based on the young person's own world constructs and beliefs. This also allows for goal focused problem solving, which has been shown to be the most effective form of problem solving (Pameijer, 2006).

1.2.6 Reinforcement and Punishment of Challenging Behaviours

When referring to a reinforcer throughout this thesis it shall refer to a stimulus which increases positive behaviours (as defined by Cooper et al, 2007). The use of contingent reinforcers for positive behaviour has been shown to be effective in altering and maintaining behaviour change (Cooper et al, 2007). Positive reinforcement (given immediately after the replacement behaviour has been enacted) encourages increased practice of replacement behaviours, increasing the use of positive behaviours over time (Cooper et al, 2007). The use of reinforcement to increase behaviours derives from behaviourism stemming from the stimulus-response explanation of behaviour by Watson in the early 1900's (Cooper et al, 2007). Behaviourism evolved and Skinner (1971) developed the definition to add types of behaviour that were determined by pre-learned consequences (operant behaviours) as well as behaviour triggered by immediate preceding stimulus (respondent behaviour). These theories developed into what is now referred to as Applied Behavioural Analysis (ABA), described as a science aimed at developing technology for improving and altering behaviours through reinforcement and punishment (consequences that decrease behaviours) (Cooper et al, 2007). Historically ABA has initiated debate in the psychological community, with those against the use of ABA claiming it aims to make a child "be still, be quiet, be docile" (Winnet & Winkler, 1972, p499). This is due to the understanding that ABA aims to stop

or increase behaviours through manipulation of consequences and therefore does not address the needs expressed through the behaviour. However, supporters of ABA would argue that when used appropriately, and where reinforcement is based on FBA (which considers internal and external antecedents and consequences for behaviour (see 1.2.3), ensuring replacement behaviours are taught to fulfil an identified function, a client is able to achieve the same function in a more socially appropriate manner (Miller, 2008).

The use of reinforcement procedures to alter behaviour has a large historical evidence base including evidence using reinforcement in the classroom to manage consequences and antecedents to effectively alter behaviours (Merrett & Wheldall, 1987). Differential reinforcement is where positive behaviours (or omission of negative behaviours) are rewarded and reinforcement is withheld for negative behaviours (Porter, 2007). There are a number of types of differential reinforcement; differential reinforcement of alternative behaviours (DRA); differential reinforcement of lower behaviour rates (DRL), differential reinforcement of incompatible behaviours (DRI); differential reinforcement of other (omission of behaviours) (DRO) (Porter, 2007; Cooper et al, 2007). A review of behavioural interventions in primary schools in the UK suggests that strategies using reward systems are effective in decreasing challenging behaviours, although additional systems are required to maintain these effects (Evans et al, 2003).

1.2.7 The Impact of Monitoring Challenging Behaviour

A number of research papers have explored the impact of pupil target setting and monitoring on challenging behaviours. These procedures may provide a way of combining the theories of reinforcement schedules (see 1.2.6) and solution focused approaches to ensure pupil participation (see 1.2.5). Johnson & White (1971) concluded, that self-observation of challenging behaviours is often used in research to measure intervention outcomes, but had rarely been considered as something that may elicit behaviour change in itself. Johnson & White (1971) aimed to research the impact of self-monitoring independently from other factors identifying this as an under researched area. The hypothesis leading to their research was based on the idea that knowledge of one's own results begins self-initiated behaviour change; this idea stems from learning research showing that knowledge about performance impacts on learning

behaviours (Johnson & White, 1971).

Johnson & White (1971) had 97 undergraduate students randomly assigned to 3 conditions: learning behaviour self-recording; dating behaviour self-recording; and control. The grades in weeks 5 and 6 suggested that the study self-recording group were involved in significantly higher levels of useful study time than the other groups. Although no difference was seen in overall grades, researchers concluded this was due to the nature of the course (a course with unusually high pass rates), suggesting that self-recording of positive behaviours can increase these. The authors assumed as self-recording can increase positive behaviours, it could also decrease negative behaviours (Johnson & White, 1971).

Maletzky (1974) conducted single case experimental design (SCED) research, focusing on whether self-recording decreases "maladaptive" behaviours. This research involved 5 participants (3 adults, 2 children (aged 9 and 11 years)) whose behaviours impacted negatively on others. All participants decreased their maladaptive behaviours when they were self-recording their behaviours. However, all behaviours increased again when they returned to baseline conditions, but a gradual "thinning" of sessions increased longevity of effects. The study, conducted by Maletzky (1974), therefore supports the assumptions made by Johnson & White (1971) that self-monitoring of behaviours can decrease negative behaviours. However, most of the participants reported wanting to change their behaviour prior to self-monitoring implementation; consequently self-monitoring may have played a supportive role to accompany the already present motivation and desire to change (Prochaska & Diclemente, 1998), rather than independently leading to behaviour change.

McNamara (1979) conducted SCED research in a selection of secondary schools where students self-recorded their own behaviours; when combined with regular feedback from staff this successfully decreased challenging behaviour. Self-monitoring combined with staff feedback allows for students to be supported with achievable mastery-based goals, which have been linked to supporting students with challenging behaviour, learning beliefs, and use of coping strategies in school (Wolters & Dougherty, 2007).

However, Salzburg (1972) (as cited in Pereira & Winton) showed that while self-recording of maths behaviours increased interest, behaviour was only altered if a reward was contingent upon good performance, suggesting self-monitoring alone may not be sufficient for behavioural change.

More recently Shabani, Wilder & Flood (2001) and Neurnberger et al (2013) conducted SCED research comparing DRO with and without self-recording in participants exhibiting challenging behaviours. Both studies showed that participant behaviour decreased with DRO, but did not decrease significantly unless the reinforcement was combined with self-monitoring (although self-monitoring of challenging behaviour was often not accurate). Neurnberger et al (2013) found that behaviour change was only maintained if students were taught to generalise their self-recording into the classroom. Studies therefore suggest that effectiveness of DRO based interventions can be increased and maintained through effective self-monitoring strategies. Few studies consider self-monitoring alone. Those that do are weakened by their generalisability to challenging behaviour (Johnson & White, 1971) or to the larger population (as studies in this area tend to be single case designs due to the heterogeneity of the population being targeted). Self-monitoring studies also tend to focus on older participants thus generalisation of these strategies to younger school populations is unclear.

1.3 Teacher Self-Efficacy and Teacher Burnout

As discussed in the introduction (see 1) there appears to be a need for research to not only explore effectiveness of training in strategies to manage challenging behaviour in relation to student outcomes (see 1.2), but also staff outcomes. Although teaching staff feel they are in need of training in challenging behaviour (Giallo & Little, 2003), little research discusses the impact of such training on staff ability to cope. This next section focuses on training in challenging behaviour theory and strategies and the potential outcomes for teacher self-efficacy and burnout outcomes of.

1.3.1 Self-Efficacy

Self-efficacy has been defined by Bandura (1977) as the conviction that one can successfully execute the behaviour required to produce desired outcomes. Teacher self-efficacy concepts have been separated into multiple constructs by some self-efficacy literature authors. Dellinger, Bobbett, Olivier & Ellett (2008) further describes self-efficacy beliefs as the factors that mediate the relationship between knowledge and the behaviours the person will execute in a given situation. Bandura (1977) reports there are 4 sources that contribute to the formation of a person's self-efficacy. These sources are mastery experiences (perceived success of previous experiences), vicarious experiences (previous experience of watching others be successful), verbal persuasion (the feedback and encouragement others give), and physiological and affective states (perceived ability to cope with situations due to physiological and affective state). Hong (2012) found that self-efficacy beliefs were a factor contributing to teaching staff leaving the profession. Kurt, Duyar & Calik (2012) suggest that individual teacher self-efficacy is impacted significantly by the collective self-efficacy and leadership styles within schools. This may be due to the way in which these factors shape the vicarious experiences and verbal persuasion aspects of self-efficacy (Bandura, 1977). However, as a person becomes more experienced in a field, their self-efficacy relating to that field becomes more difficult to influence through external forces, as their own experiences have a greater impact (Palmer et al, 2010).

Ware & Kitsantis (2007) conducted research showing that the higher self-efficacy of teachers, the higher their professional commitment to teaching, thus highlighting its

importance and relevance to the teaching profession. Lee, Cawthorn & Dawson (2013) report that there are a number of research papers demonstrating that student outcomes and teacher self-efficacy are positively correlated. This has been shown to be subject specific, with teachers with lower knowledge of science and maths having low confidence in their ability to teach these subjects (Clerici, 2008). Training has been shown to be able to increase self-efficacy in teachers, suggesting that commitment to teaching may be reinforced through increased training and knowledge (Jennett, Harris & Mesibov, 2003). Gibbs & Powell (2012) conducted a factor analysis using data from 197 teachers from 31 primary schools and found self-efficacy beliefs to be separated into 3 distinct factors; classroom management, children's engagement, and instructional strategies. A review of literature suggests there is a significant gap in literature examining self-efficacy in relation to challenging behaviour.

1.3.2 Teacher Burnout

Teacher burnout has been defined as the point at which a person is no longer able to cope with chronic stress (Jennett et al, 2003). Maslach & Jackson (1981) described teacher burnout as a syndrome with "three dimensions": emotional exhaustion; depersonalisation; and reduced feelings of personal accomplishment. Farber (1991) (as cited in Jennet, Harris & Mesibov, 2003) described teachers who are "burned out" as emotionally exhausted, displaying attitudes that depersonalise students, and displaying low levels of personal accomplishment from their own work. A review suggests that this low level of personal accomplishment predicts teacher burnout levels (Hastings & Bham, 2003).

Teachers deciding to leave the profession have been shown to have increased levels of burnout and stress (Hong, 2012). Federici & Skaalvik (2012) found that teacher burnout is positively related to the desire to quit, and negatively related to job satisfaction. Teacher burnout ratings were also found to be negatively correlated to self-efficacy levels (Federicki & Skaalvik, 2012). A review of research by Hastings & Bham (2003) found that student challenging behaviour levels are significantly and positively related to teacher burnout levels. Mitchell & Hastings (2001) showed that school staff's negative emotional reactions to challenging behaviour are related to their levels of

burnout, therefore highlighting the importance of reducing challenging behaviour, possibly through increasing teacher knowledge and skill set. Eyged & Short (2006) demonstrated that teachers who are less sure of how to manage students with challenging behaviours are more likely to have higher levels of teacher burnout. In a study of 558 secondary schools, staff perceived self-efficacy in classroom management was shown to be a guiding construct in teacher burnout (Brouwers & Tomic, 1999).

1.3.3 The Relationships Between Teacher Self-Efficacy, Teacher Burnout, and Student Academic and Behavioural Outcomes

Eyged & Short (2006) used vignette methodology to explore teacher self-efficacy and burnout ratings and their relationship with special educational needs (SEN) referral patterns. While a significant relationship was found between teacher burnout and challenging behaviour, and decision making, this was not the case with self-efficacy ratings. However, self-efficacy and teacher burnout ratings were shown to have a strong significant inverse correlation (as self-efficacy decreased, burnout increased). This suggests that while both teacher burnout and self-efficacy may play different and independent roles in the management of challenging behaviour they are related constructs. For this reason, as well as the notable overlap in teacher self-efficacy and teacher burnout research, these concepts are discussed under one heading. Although they are not necessarily assumed to be measuring the same concept, the research suggests they are related.

High levels of stress have been shown to be connected with increased desire to leave the teaching profession (Weisberg & Sagie, 1999). Increased stress has also been linked to low teacher self-esteem, which is connected to the measure of teacher burnout (Jennet et al, 2003). This demonstrates the importance of reducing stress levels in schools. It may be that increased teacher self-efficacy and lower levels of teacher burnout reduce stress, thus increasing job satisfaction (Ware & Kitsanas, 2007). Increased job satisfaction from teachers may be influential in increasing willingness to implement and persist with differing teaching strategies associated with high teacher self-efficacy (Klinger, Ahwee & Pilonieta, 2003), leading to increased positive outcomes for students.

Teacher self-efficacy levels are positively associated with student academic outcomes (Kurt et al, 2012). A review of research (Tschannen-Moran, Woolfok Hoy & Hoy, 1998) suggests that self-efficacy impacts a number of classroom outcomes, especially teacher ability and classroom ethos. High levels of instructional self-efficacy were found to lead to increased time spent focusing on classroom activities, a stronger commitment to teaching, and teachers who were more willing to try out new teaching strategies and interventions. High self-efficacy levels were also found to lead to positive student outcomes such as increased help for students with difficulties and increased student motivation and self-regulation (Tschannen-Moran et al, 1998). Motivation and self-regulation have been found to be important skills for achieving both academically and socially in educational settings (Dawson & Guare, 2010). This research suggests that by increasing teacher self-efficacy, teachers become more able to support skills crucial to educational success in students.

There have been a number of studies focusing on teacher self-efficacy, teacher burnout and their relationship to challenging behaviour. Research has shown that challenging behaviour can lead to teacher burnout (Hastings & Bham, 2003) and that burnout is linked to an increase in physical, mental, and personal difficulties for teachers (Hastings & Bham, 2003). Specifically challenging behaviour that teachers interpret as "disrespectful" and low "sociability" was linked to an increase in teacher emotional exhaustion and depersonalisation (Hastings & Bham, 2003). Mitchell & Hastings (2001) state that teacher burnout levels are predicted by their negative emotional reactions to challenging behaviour. This suggests that teachers with higher burnout levels will be less effective in managing students with challenging behaviour due to negative emotional responses. Robertson & Dunsmuir (2013) found that high self-efficacy and low levels of negative comments about student behaviour predicted an increase in pupil on-task behaviour. A similar situation is likely to occur when teachers with low self-efficacy are dealing with challenging behaviours, demonstrated through a questionnaire study of 70 teachers of students with autism spectrum condition (ASC) (Jennet et al, 2003). This study showed that staff with low self-efficacy about ability to manage behaviours showed increased fear, anxiety, depression, and angry reactions with students. Research such as that of Gibson & Dembo (1984) supports this; showing that classroom behaviour is better when teacher self-efficacy is higher. However, this is

correlational research so cannot determine causality. Yet when considered as part of the larger field the research indicates that teacher self-efficacy does have a causal impact on challenging behaviour of students (Kurt et al, 2012).

Research conducted by Hastings & Brown (2002) showed that higher levels of behavioural knowledge decreased teacher depression, anger, and emotional reactions when faced with challenging behaviour in the classroom. There have been mixed reports on whether increased knowledge around a topic increases self-efficacy for teachers. Officer's (2012) research contradicts Hastings & Browns' (2002) as it found no relationship between middle school teachers' training in adolescent mental health and teacher self-efficacy in managing students with emotional or behavioural needs. This could be because of the assumed rather than explicit link between mental health issues and behaviour management strategies; teachers may need more specific training on how to implement theory into practice in order to improve their self-efficacy. This may relate to the findings showing that specific behaviour management techniques and training increase knowledge and teacher self-efficacy in regards to managing behaviours (Crone et al, 2007). Crone et al (2007) found that through specific intensive teacher training in FBA, teacher understanding of challenging behaviours, self-esteem, and ability to implement strategies all increased, student challenging behaviour also decreased. However, de Courcey-Bower (2011) found that explicit links between theory and training (in FBA) were missed by teachers, decreasing programme effectiveness. This suggests training that aims to improve teacher self-efficacy and burnout (as emotional reactions are decreased) through increasing knowledge, should make explicit links between theory and practice (Lang, Sigafoos, Lanconi, Didden & Rispoli, 2010). This, combined with the information that challenging behaviour is an area where most teachers feel they are lacking in training and knowledge (Hemmeter, Fox, Jack & Broyles, 2007), suggests training is necessary to increase knowledge around challenging behaviour that links directly to a behaviour plan. Brinson & Steiner (2007) showed that by altering teacher self-efficacy through training, student learning can be positively impacted. This highlights training as an area warranting further detailed exploration as a potential method for increasing staff well-being (through increasing perceived teacher competence, therefore decreasing teacher burnout (Hastings & Bham, 2003)) and positive student outcomes (through decreased challenging behaviour, (DfE, 2012a)).

1.4 Systematic Review of Research Exploring Interventions that Impact Teacher Self-Efficacy and Teacher Burnout

1.4.1 Background

There are a number of studies considering the impact of training on self-efficacy and teacher burnout. This systematic review aims to synthesise research that explores interventions aimed at increasing teacher self-efficacy and decreasing teacher burnout through teacher training or other interventions. It has been discussed above (see 1.3.3) that teacher self-efficacy and burnout can impact a variety of factors including willingness to engage with teaching strategies (Klingner et al, 2003) and teacher attrition (Weisberg & Sagie, 1999). This thesis aims to explore the impact of training on teacher self-efficacy and burnout. This systematic review will therefore explore research that has aimed to impact teacher self-efficacy and burnout through implementation of strategies and interventions directly with the teachers. As the research area is relatively limited, the review considers all aspects of teacher self-efficacy, not restricting itself to self-efficacy regarding only challenging behaviour. This will allow the consideration of factors that are able to affect self-efficacy and burnout in teachers in general, although it will be noted which studies focus on these concepts in relation to challenging behaviour, given the scope of this thesis. This review aims to synthesise research evaluating interventions and training in educational settings to determine whether there is a conclusive evidence base suggesting training will increase teacher self-efficacy and decrease burnout.

1.4.2 Statement of Review Purpose

To review research that assesses the impact of training and evaluations with staff in educational settings on their own self-efficacy and burnout.

*1.4.3 **Method***

Criteria for Considering Studies for this Review

Types of Studies

The types of articles considered for this review are peer reviewed published articles where the full text is available. The research must consider teacher outcomes of self-efficacy and burnout. Teachers must be in educational settings teaching children between ages 2-19 years old. Due to the nature of the research area this review does not stipulate specific methodology. The research must involve the implementation of an intervention that directly or indirectly aims to alter self-efficacy and / or teacher burnout levels. Methods of data collection may include quantitative and qualitative methods including observations and ratings / questionnaires / interviews of key persons involved.

Types of Participants

Participants in the research need to include teachers or teaching staff who work in school settings with young people aged 2–19 years, as these participants are the focus of the educational services supplying the training within this thesis.

Types of Interventions

This review includes studies using / evaluating any type of intervention and / or training with teachers that is aimed at directly or indirectly impacting teacher self-efficacy or burnout. This can be compared with no treatment / alternative treatment, with the participants' own baseline measures or across participants.

Types of Outcome Measures

Primary or secondary outcomes of the studies must include teacher self-efficacy and / or teacher burnout measures. These may also incorporate measures that have been shown to make up these concepts (e.g. teacher stress). Secondary outcomes that may also provide useful information may include student outcomes due to implementation of information / strategies gained as part of the intervention by the teachers. Measures can be qualitative and / or quantitative and can take the form of observation of measurable

behaviours, interviews, questionnaires, standardised or criterion referenced tests, or any measure that has evidence of its reliability.

Search Methods

Studies were identified from the following databases:

PsycINFO (1806 to July Week 1 2012) (last accessed 23-07-2013)

Web of Knowledge (1899 to 2012), (last accessed 23-07-2013)

Cochrane Central Register of Controlled Trials (Central: July 2012) (last accessed 23-07-2013)

The group terms used to identify articles were

1. teacher
2. self-efficacy
3. burnout
4. training

Terms were adapted to suit the fields of each of the databases and searched for either as part of topic, keywords, or abstract. Searches of 1, 2, 3, and 4 were combined with "AND" and the terms within groups were combined with "OR" or searched for separately according to the search fields (see Appendix 1 for details of database searches and results). Searches were also conducted using "AND" with combinations of search terms (1, 2, and 4) and (1, 3, and 4). According to database where necessary, research results were refined by peer-reviewed published article, full article, and school. (See Appendix 1 for full search details).

Limitations of Systematic Review

In order to ensure this review is systematic parameters must be placed around the search criteria and processes. While ensuring a systematic and fair search this can also create weaknesses and potentially reduce the number of studies found. By using electronic databases only it is likely that some research may not be included in the review, due to age of research or publication type, however this is also able to minimise risk of articles

being selected due to arbitrary factors such as copies made available to certain libraries etc. Searching for only published and peer reviewed articles allows for a level of quality assurance, however also risks the systematic review being influenced by publication bias (the tendency for articles with positive findings to be published over those with zero or negative findings (Müller et al, 2013)). The search terms used within this systematic search are limited in their scope, which minimises irrelevant articles being found as a result of the search, but also risks some relevant articles being missed due to use of related terms. In order to try to minimise this, the functions "explode" and "related terms" were selected during all searches conducted.

1.4.4 Data Collection and Analysis

Selection of Studies

From the search results, studies were assessed by their title. They were excluded if the title showed a lack of implementation / evaluation of an intervention, did not focus on the target population (teachers of school age pupils), or outcome measures were unrelated to the key outcomes. Abstracts were studied to acquire further details about meeting the search criteria – duplicates of any articles and studies not meeting search criteria were removed. Where the abstract was ambiguous the whole text was assessed for meeting inclusion criteria. Where full texts were not available studies were excluded. Full texts were read and included in the articles reviewed if they met the criteria for intervention and outcome measures as stated above.

Data Extraction and Management

This systematic review is qualitative in nature and therefore quantitative data was not collected in its original form for analysis. However, results are discussed according to their numerical representation in the research articles. Information from each study was extracted by the author and categorised according to a set of criteria stipulated in the Cochrane Handbook for Systematic Reviews of Interventions (Higgens et al, 2008). The information was extracted ensuring that these criteria were considered for each article prior to the review and amalgamation of data from all articles:

Participant demographics*, Intervention, Duration*, Outcomes*, Design*, Number of groups / Comparisons, Results, Outcome Measures (when and how collected), Blinding of assessors, Participant selection and missing participants, Investigation question*, Author conclusions, Potential limitations, Reliability.

(* items relating to inclusion criteria)

Assessment of Risk of Bias

All articles were assessed for bias in relation to selection, attrition, detection, reporting, and performance (Appendix 2). This table has been taken from Higgens & Altman (2008) and simplified to meet with the aims of this systematic review.

*1.4.5 **Results***

Search Results

The total number of citations given through the database searches was 144, 140 studies remained after removal of duplicate citations. Titles were scrutinised for relevance of topic and intervention (see Appendix 1 for study exclusion information) leading to a total of 56 citations. From these 56 abstracts, a total of 27 citations met criteria for full text analysis. During full text analysis a further 10 were excluded as full texts were not available to the author, resulting in 17 texts for full text analysis. A further 4 citations were excluded from this review as they did not meet inclusion criteria (see Appendix 2). The reasons for their exclusion included a lack of intervention implementation / evaluation (Chan, 2005; Clerici, 2008), and results not having yet been published (Ford et al, 2012) (see Appendix 2).

Setting of Research

The studies included in this systematic review were implemented in a variety of different settings where teachers and trainee teachers attended courses / work. The research was conducted in a number of countries including Israel, America, Australia, and the UK. Studies set in mainstream schools are more common within this literature review. 9 of the studies were set in mainstream schools. 7 of the 13 studies were conducted with staff from multiple primary schools; Britten & Lai (1998) conducted research in 6 mainstream elementary schools, Palmer (2010) in 12 elementary schools, Shechtman, Levy & Leichtentritt (2005) recruited staff from 97 elementary schools, Stoiber & Gettinger (2011) conducted research with pre-kindergarten, kindergarten, and first grade teachers. Revital (2009) and Telljohann, Everett, Durgin & Price (1996) recruited teachers from multiple elementary schools. Lee et al (2013) was the only study to include staff from primary and secondary schools, recruiting participant teachers from 12 elementary schools, and 18 secondary schools. Staff recruited only from individual secondary schools took part in 2 studies; Newman-Carlson & Horne (2004) conducted research with staff in an individual middle school, and Kaspereen's (2012) research involved staff from a high school. Jennet et al's (2003) research is the only study in this review to work with teachers in special education settings, where staff implemented either ABA or TEACCH programmes with young people with ASC. The

review includes 3 studies where the interventions were based in non-school settings. These included participants from a Masters programme (Hall, Hall & Abaci, 1977), a teacher training programme (Liaw, 2009), and a training programme for music teachers (Hargreaves, Purves & Graham, 2007).

Designs of Research

Almost all the studies in this review are mixed method with the exclusion of Jennet et al (2003); Newman-Carlson & Horne (2004); and Kaspereen (2012) which used only quantitative data. All studies used some form of self-report measures. While all studies revolved around evaluation of specific interventions, Britten & Lai (1998), Revital (2009), and Liaw (2009) evaluated an intervention that was already in place for all participants and therefore only have during and post data without a comparison group. Britten & Lai (1998) compared differences between amount of training within their experimental group to determine the impact of the intervention. In Jennet et al's (2003) study, staff already implementing one of two interventions were compared to one another on various outcome measures. All other studies included have either comparison or control groups (Revital, 2009; Hargreaves et al, 2007; Shechtman et al, 2005), pre- and post- comparisons (Palmer, 2010; Lee et al, 2013) or both (Hall et al, 1997; Stoiber & Gettinger, 2011; Telljohann et al, 1996; Kaspereen, 2012).

Of the studies 3 were entirely based on (or involved a hypothesis that used) randomised control trials (RCT) (Hall et al, 1977; Stoiber & Gettinger, 2011; Kaspereen, 2012). 2 studies used matched pairs / group designs (Hargreaves et al, 2007; Telljohann et al, 1996). Other studies were either grouped according to convenience / qualities that determined analysis, or were single group designs. No single case designs (SCED) were used, although Stoiber & Gettinger (2011) used a quasi-experimental design involving a secondary hypothesis relating to student outcomes that used a SCED for some target students. Other studies used retrospective analyses (Jennet et al, 2003; Britten & Lai, 1988) as well as exploratory factor analysis (Britten & Lai, 1988), and longitudinal data (Hargreaves et al, 2007).

Target Populations and Participant Selection

The participants focused on for the purpose of this review were teachers and trainee teachers of school age children. Although 1 study (Stoiber & Gettinger, 2011) also investigated impacts on students themselves, as these student-oriented results are not relevant to the purpose of this review, they are not discussed within this review.

All participants in the studies taught in schools, with the exception of the participants of Hargreaves et al (2007) who compared undergraduate music students (control group) to trainee music teachers (experimental group). For the majority of the studies participant selection was completed through request for volunteers from opportunity samples based on setting and / or location. This was not the case for Hargreaves et al (2007) and Liaw (2009) where participation was compulsory as part of a course. The study with the most participants is Shechtman et al (2005) with 360 5th and 6th grade teachers. The number of participants in the other studies range from 12 (Palmer, 2010) to 324 teachers (Britten & Lai, 1998).

Types of Intervention

The studies in this review have incorporated interventions and training in various subjects and topics to explore the impact of this on teacher self-efficacy and / or teacher burnout. The amount of information given regarding the specific training within each study varies significantly, often depending on the research's main purpose. Jennet et al (2003) compared 2 types of intervention (ABA and TEACCH) and did not themselves provide the training for the participants. Therefore despite knowing that all participants received specific training on 1 of the 2 approaches, it is not clear how extensive the training was. Britten & Lai (1998) did not provide the training in nutrition teaching, but evaluated impact based on retrospective teacher reports concerning the amount of training received. 2 of the studies provide minimal details about the training used, as the focus of their research is the evaluation of standard training courses, these were Liaw (2009) who evaluated a teacher preparation module of a teacher training programme, and Hargreaves et al (2007) who evaluated impact on values of a music teaching teacher training programme. Revital (2009) did not implement the intervention evaluated, but the training was familiar to the researchers as it was developed and provided in a

specific school district. This training in school violence prevention required teachers to attend weekly for 2 years, (participants completed this to varying degrees). This training aimed to provide theoretical knowledge around violence issues, exploration of various violence management strategies, and enhance skills in managing violent situations.

The final 8 studies involved an intervention / training with the researcher's involvement. Kaspereen (2012) was the only study to not involve training, providing relaxation therapy for 30-45 minutes per week for 4 weeks with the specific aim of decreasing teacher stress (burnout factors). Palmer (2010) delivered an intervention that specifically targeted self-efficacy through training. Palmer (2010) developed a 6 week training course aimed to enhance teacher self-efficacy in science teaching, incorporating cognitive mastery, enactive mastery, modelling and verbal persuasion with opportunities to observe and practice skills.

Training that was not specifically designed to improve self-efficacy and burnout levels was implemented in the remaining studies. Newman-Carlson & Horne (2004) provided 3 training sessions in bullying prevention. Lee et al (2013) provided an initial 6 hour training session and then regular mentoring sessions as well as small group training sessions for 1 academic semester in the using drama techniques in the classroom. Hall et al (1977) evaluated the impact of a 2 year human relations Masters (3 hours a week for 3 10-week terms). Telljohann et al (1996) evaluated impact of "Project Healthy Kids"; a 30 hour training programme focusing on teaching health education. Stoiber & Gettinger (2011) developed a 15 hour training programme in FBA and positive behaviour support which including theory and joint planning. This is the only research derived from the search that uses training in behaviour management theory and techniques.

Measures and Reliability

All studies in this review use quantitative self-report measures (usually in the form of questionnaires) risking effects of response-bias, desirability effects, and subjective data. 8 of the studies used only self-report questionnaires (Jennet et al, 2003; Britten & Lai, 1988; Newman-Carlson & Horne, 2004; Revital, 2009; Hargreaves et al, 2007; Shechtman et al, 2007; Telljohan et al, 1996; Kaspereen, 2012; Lee et al, 2013). 3 of the

studies used interviews, therefore relying on subjective interpretations and inter-rater reliability, and again risking social desirability effects, (Palmer, 2010; Hall et al, 1977; Liaw, 2009). Only Stoiber & Gettinger (2011) used a data collection method other than self-report questionnaires / forms and interviews. They combined data from self-report measures with data from observation (risking observer bias as observers were not blind to condition).

5 of the studies developed their own questionnaires with differing levels of reliability and validity testing prior to use (Britten & Lai, 1998; Lee et al, 2013; Revital, 2009; Hargreaves et al, 2007; Telljohann et al, 1966). 5 studies used the teacher self-efficacy scale (Jennet et al, 2003; Newman-Carlson & Horne, 2004; Lee et al, 2013; Liaw, 2009 (adapted); Shechtman et al, 2005) and 2 studies used the Maslach Burnout Inventory (Jennet et al, 2003; Hall et al, 1977). 7 of the studies used other pre-devised scales relating to either their specific intervention or outcome measures (Jennet et al, 2003; Newman-Carlson & Horne, 2004; Palmer, 2010; Hall et al 1977; Stoiber & Gettinger, 2011; Shechtman et al, 2005; Kaspereen, 2012). None of the studies used assessors blind to the experimental aims or conditions. However, most used quantitative self-report questionnaires which should minimise impact of non-blinded assessors.

Intervention Results

A number of these studies included research questions irrelevant to this review's purpose. For this reason, only results relevant to self-efficacy and burnout will be reported. Jennet et al (2003) compared 2 groups of teachers using different teaching strategies (ABA/TEACCH) and how commitment to teaching philosophy impacted self-efficacy and burnout. They found commitment to teaching was significantly positively correlated with personal self-efficacy scores ($p<0.05$). Teaching commitment was positively correlated with general self-efficacy for the ABA group ($p<0.001$) and negatively correlated with burnout scores for the TEACCH group ($p<0.05$) only.

Britten & Lai (1998) found that nutritional knowledge was significantly correlated with self-efficacy ($p<0.05$), and with time spent teaching and in training ($p<0.05$). By fitting these to linear models they found that the data fitted the path co-efficient: self efficacy

leads to increased time spent training, which increases self-efficacy and knowledge, increasing self-efficacy ($p<0.05$). The data also fitted a second model with the same level of significance: self-efficacy impacts time spent training and time spent in training impacts knowledge, impacting training.

Newman-Carlson & Horne (2004) found that training in bullying prevention for 6th, 7th and 8th grade teachers significantly increased teacher knowledge, specific areas of self-efficacy in behaviour management, but not general self-efficacy compared to pre- or control group measures. Palmer (2010) found that training in science teaching significantly increased self-efficacy through cognitive mastery (effect size = 1.24). Self-efficacy was also significantly increased compared to pre-training scores immediately post and 2 years after the training ($p=0.002$), no differences were found between post and delayed measures of self-efficacy.

Lee et al (2013) found that elementary school teachers had higher self-efficacy than secondary school teachers ($p<0.05$). Training in drama-based teaching techniques was not found to alter self-efficacy. However, in secondary school teachers, as self-efficacy increased the amount of conceptual change also increased ($p<0.01$). Stoiber & Gettinger (2011) found, that training in FBA and positive behaviour support led to an increase of reported feelings of competence and self-efficacy ($P<0.01$) as well as an increased observed utilisation of skills ($p<0.001$).

Revital (2009) reported that as the amount of training in prevention of school violence increased so did teacher perceived outcome self-efficacy in managing violent situations ($p<0.001$), but this did not alter personal or general self-efficacy. Perceived outcome self-efficacy was significantly higher as experience increased. Liaw (2009) found that a teacher preparation programme increased trainee teachers' self-efficacy to motivate students, manage a classroom and select appropriate learning materials. However, self-efficacy to manage environmental factors decreased. Shechtman et al (2005) found a significant positive correlation between teacher self-efficacy and supervisor support, and clarity of rules. 2 years of life skills training resulted in significantly higher self-esteem for teachers than 1 year skills with no training, with class size and school SES

acting as mediators.

Telljohann et al (1996) found a significant increase in self-efficacy ($p=0.03$), outcome expectations ($p=0.03$), and hours teaching health education ($p=0.01$), in teachers who were a part of the "Healthy Kids Project" compared to controls. The changes within the experimental group were significant for efficacy expectation ($P<0.001$), outcome expectation ($p<0.001$), hours spent teaching health education ($p=0.002$), and time and effort spent on health education ($p<0.001$).

Hargreaves et al (2007) was the only study not to find an effect of training / intervention on self-efficacy or burnout, although there was some difference in attitudes towards education between undergraduate music students and trainee teacher music students.

Hall et al (1977) found that training in humanistic classroom management resulted in a shift to humanistic ideology ($p<0.01$), increased teacher sense of personal accomplishment ($p<0.05$), and reduced emotional exhaustion ($p<0.01$) compared to pre- and control measures. Kaspereen (2012) also considered burnout as a dependent variable and found that teachers who underwent relaxation therapy reported significantly lower perceived stress ($p<0.001$), work stress ($p<0.001$), and increased life satisfaction ($p=0.007$) than controls.

In summary, all studies investigating impact on self-efficacy found positive significant effects for some aspects of self-efficacy (but not all subscales) other than Hargreaves et al (2007). Both studies considering teacher burnout (Hall et al, 1977; Kaspereen, 2012) found training / intervention led to a significant decrease in some factors of burnout.

Potential Limitations of Methodologies

All studies used self-report measures and are therefore open to social desirability factors, and subjectivity difficulties. However, as the studies are focused on perceived self-efficacy and perceptions of internal factors it would be difficult and potentially impossible to ascertain these results through any other method. Some studies have

included interviews and observations. Interviews are also vulnerable to the same weaknesses as self-report measures, and observations are vulnerable to observation bias. If results were to be based only on observation data researchers would need to assume a connection between action and internal concepts to make results relevant to self-efficacy and / or burnout. Studies are especially vulnerable to social desirability effects where the researcher is the same person who delivers the training / intervention or where participants are aware of the research aims (Kaspereen, 2012).

Studies that sent questionnaires out for return are subject to selection bias, as are studies using volunteers. These selection procedures were used in all but 2 studies (Hargreaves et al, 2007; Hall et al, 1977), in which participation was compulsory. These are likely to be more vulnerable to response bias as research participation was part of their course. A number of the studies do not explicitly account for extraneous variables such as teacher experience, SES, and position in the school hierarchy, although group methodologies with a number of individual differences within groups aim to counteract this somewhat.

Of the studies, 3 used an RCT design (Hall et al, 1977; Stoiber & Gettinger, 2011; Kaspereen, 2012). 1 study has a purely exploratory design (Britten & Lai, 1998) subjecting it to scrutiny due to its lack of specific data, low control over variables, and reliance on researcher interpretations. However, the weight of evidence model (Gough, 2007) states that studies should be clear, transparent, and fit for purpose. Therefore as Britten & Lai's (1998) purposes required factor analysis to hypothesise about models this seems appropriate as a method for their particular question. All studies used group designs, those without control groups were more susceptible to maturity effects. Some studies did not include pre-intervention or comparison data (Britten & Lai, 1998; Revital, 2009; Liaw, 2009), potentially decreasing validity as there may be unaccounted for variables relating to outcome measures prior to intervention.

1.4.6 Discussion

Most of the studies from this literature review had self-efficacy as an outcome rather than burnout. Self-efficacy was explored in 11 studies, and burnout was explored by 2 studies. This suggests a decreased interest in teacher burnout despite literature linking the concepts (Hong, 2012). All studies recruited their participants either voluntarily or through compulsory course requirements. Both present their own difficulties, as volunteer participants may have different characteristics impacting outcomes, while those who must take part due to course requirements may be more likely to answer with what they feel are socially desirable responses. However, as a whole, the research showed that in a variety of designs, participant selection procedures, and settings, training tended to improve self-efficacy (although usually specific aspects of self-efficacy rather than general self-efficacy) and decrease burnout factors.

Only 1 study did not support any of its hypotheses in regards to teacher burnout and / or self-efficacy. This was Hargreaves et al (2007) where trainee music teachers were compared with undergraduate music students. The training was part of the standard music teacher training course and self-efficacy relating to music was not altered in trainee teachers or undergraduates. This may be due to the specificity of the subject. The comparison of undergraduate students may also impact effects as they are not comparable in terms of teaching experiences and it may be that measures considering teaching self-efficacy are not appropriate for non-teaching undergraduate students. Literature suggests that teaching music requires different and specific skills compared to other teaching areas (Hargreaves et al 2007). This may imply that this research is not generalizable to other teacher self-efficacy literature in exploring alternative subjects.

Other studies also considered subject specific teaching (Britten & Lai, 1998 (nutrition); Telljohann et al, 1996 (health); Palmer, 2010 (science)) and found an impact of training on self-efficacy. This may suggest that there are aspects of either the training, research design, or subject that led to change. These studies used a variety of designs including control and pre- and post-measures as well as different intensity rates of training, but none of the studies consider comparison groups. This may suggest impacts were due to the implementation of an intervention rather than the intervention specifics or type. An

alternative interpretation may be that any increase in knowledge through training will improve teacher self-efficacy ratings. Hargreaves et al (2007) may not have found an impact as the trainee teachers received the training as part of their standard course and therefore may have already experienced changes by the time they were being researched (final term of the course).

Jennet et al (2003) researched teachers who had already received training, meaning specific training details could not be given. Jennet (2003) found teachers with increased commitment to teaching methods had increased self-efficacy (professional only) and suggested that different teaching strategies impacted different aspects of self-efficacy (e.g. general self-efficacy was only correlated with commitment to teaching for the ABA group). It is not, however, possible to determine the type or intensity of training between the 2 teaching strategies, as the teaching strategies are very specific. It may also be that as teachers are likely to have chosen these specific teaching methods due to an already existing commitment to the philosophies, therefore their experiences may not be representative of teachers in more general teaching environments. Given the scope of this review, and thesis, training details are important to determine what is required to produce outcomes in self-efficacy and burnout. The majority of the studies in this review used intensive long-term training interventions, possibly decreasing applicability into schools given the current economic climate where resources are being cut (Sellgren, 2013). Considered individually many of the studies are not generaliseable, especially studies with smaller participant numbers (Palmer, 2010), or in specific settings (Jennet et al, 2003). A number of studies use large sample numbers and compare groups of teachers, it may be that teacher groups are very idiosyncratic. This might be especially important to consider in further research as research has shown differences in target outcomes due to experience and hierarchical factors (Penrose et al, 2007).

When combining research from the field, the studies in this review incorporate a number of different settings including mainstream elementary, secondary, and special schools. This implies that generally increased training increases self-efficacy / decreases burnout in teachers (although this may be subject-specific). When assessing the review

in relation to the purpose of this thesis and considering potential idiosyncratic factors of different training areas only 2 studies in this review explore target outcomes in relation to challenging behaviour (Revital, 2009; Stoiber & Gettinger, 2011). Both used intensive training programmes, and neither explored specific factors of the training that impacted self-efficacy measures. It may be that training increases knowledge, therefore increasing self-efficacy (Britten & Lai, 1998), possibly explaining the positive effects found in multiple varied interventions.

In light of the reviewed research it appears that training and / or interventions that aim to increase knowledge around theory and implementation of strategies in a particular teaching area can increase self-esteem and / or decrease burnout. Limitations in design and generalisability leave the research base unclear as to individual factors required within training to impact self-efficacy and burnout. The review suggests there is likely to be an impact of training on teacher self-efficacy and burnout, although this may be different according to subject matter. Only 2 studies considered the impact of training in challenging behaviour on self-efficacy and burnout. This highlights a distinct lack of research in this field, despite a significant number of studies showing the importance of challenging behaviour training and knowledge on teacher outcomes (see 1 and 1.3.3).

2 Research Questions

The literature review shows that challenging behaviour is an increasing difficulty for teachers and pupils in schools (see 1), and further discusses psychological theories that guide challenging behaviour management (see 1.2). An intervention has been developed aimed at decreasing challenging behaviour in primary school students, incorporating theories discussed in the literature review (FBA, SFP, pupil led target setting, and reinforcement (see 1.2)). This was developed by a specialist teacher at a Behaviour Enhanced Mainstream Service (BEMS) in the area where the research was conducted. Prior to this research the distribution of the intervention was limited by the individual case style of working of the BEMS, and there was no record of any formal evaluation.

The literature (see 1.2) has identified challenging behaviour as an area where teachers require increased training. The systematic review presented in section 1.4 suggests that training can increase teacher self-efficacy and decrease teacher burnout, which impact significantly on teacher and student experiences (Hastings & Bham, 2003). This research will aim to investigate the impact of training in the intervention, including information about the theories and principles on which it is based. The impact of this training on teacher self-efficacy and burnout will also be explored, as these can impact willingness to implement strategies (Klinger et al, 2003). This research will aim to evaluate training around theory and implementation of a behaviour plan to staff, and the outcomes of this on student challenging behaviour, and staff self-efficacy and burnout.

The BEMS developed an intervention using previously discussed theories of psychology (see 1). This thesis aims to evaluate the effects of intervention and accompanying staff training on staff outcomes (through an RCT design) and child outcomes (through a SCED). The research questions that will be explored are:

1) In a sample of rural primary schools can training based on a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms:

a) lead to use of the behaviour plan?

b) increase overall staff self-efficacy in behaviour management?

c) decrease teacher burnout?

2) Does use of a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil lead target setting, and the use of regular feedback mechanisms decrease challenging behaviour in students?

3 Research Context

The research was undertaken in a large county (approximately 600,000 residents (Office for National Statistics, 2011)) in the North of England covering a combination of urban and rural areas. This research was undertaken in rural and urban schools within the county in collaboration with a teacher from the BEMS. The county's Children and Young Peoples Service are currently undergoing a re-structure which may lead to a reduced number of hours specialised services can offer schools. This is not an unusual situation in the current economic climate. The recent funding cuts in education (Sellgren, 2013) have led professionals working in this field to raise concerns about the availability of specialist services and staff to provide support with managing challenging behaviour (Buie, 2009). This highlights the need to ensure that schools are equipped to handle challenging behaviour, through ensuring effectiveness and existence of effective support for staff and students, in order to minimise the need for referrals.

Key stakeholders of the research project include the BEMS teacher who developed the intervention and the Council where the research is undertaken. This research was conducted under the supervision of Nottingham University Doctorate of Applied Educational Psychology (DAppEdPsy). (See 4.2.3 for details of stakeholder involvement).

4 Methodology

4.1 Introduction

Research is defined in Mertens (2005) as a process of "systematic enquiry". This process aims to predict or control phenomenon through collection, analysis, and interpretation of data gathered within the context the phenomenon occurs. Mertens (2005) argues that researchers conducting research within the real world are required to have an awareness of the paradigms they work in, as well as an understanding of the influence their beliefs and values may have over the research they conduct. However, other researchers argue that research should be guided by the question rather than preference or affiliation with certain paradigms or epistemology (Bryman, 2006). This chapter will aim to explore epistemological and ontological assumptions, leading to detailed discussion of paradigms and a rationale for the epistemologies and designs used within this research (see 4.3). Limitations of a number of designs will be discussed which will guide the formulation of the research plan. This will be followed by details of the current research procedure, intervention, measures, and participants.

4.2 The Present Research

4.2.1 Real World Research

The current research is based in educational settings. Academic research tends to focus on developing an academic discipline, while real world research usually focuses on social issues and problems that directly impact on people's lives within their natural environments (Robson, 2011). Research in real world settings is often open to a number of confounding variables relating to environmental changes (Robson, 2011). Research in educational settings is at risk of ethical issues (see 4.11) of working with and gaining consent from children and young people, potentially made more complex by the addition of special educational needs (Loveridge, 2010). Research based in educational settings is known for a number of implementation difficulties, such as maintenance of staff commitment, recording subtle effects of interventions, and dealing with school change (Maruyama & Deno, 1992). However, research in educational settings is seen as crucial for understanding and improving education which, when collated effectively and efficiently, can significantly impact educational policy, improving lives of children and young people (Ozga, 2000).

4.2.2 Research Questions

The BEMS developed an intervention to support staff to manage challenging behaviour in the classroom. The intervention is based on psychological theories relating to solution focused questioning and behaviour analysis and includes weekly monitoring and rewards, with pupil led targets agreed upon through staff-pupil discussion. This thesis will aim to explore its use both in relation to staff and student outcomes. The research questions that will be explored are:

1) In a sample of rural primary schools can training based on a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms:

a) Lead to use of the behaviour plan?

b) Increase overall staff self-efficacy in behaviour management?

c) Decrease teacher burnout?

2) Does use of a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms decrease challenging behaviour in students?

4.2.3 Key Stakeholders

This research has a number of stakeholders. The thesis is undertaken as part of the DAppEdPsy course at Nottingham University. Nottingham University has supported the author to conduct a research project of an acceptable ethical and scientific standard for submission towards the DAppEdPsy course. Parameters set by the University state that the thesis must contribute to the evidence base for educational practice with due consideration to priorities of the local service where the research will be conducted. Another key stakeholder in the research is the county council where the research is based. The research is conducted with schools and services from the county council and the research is aimed to contribute positively to the Children and Young Peoples Department, particularly the Educational Psychology Service (EPS) and schools in the area.

Another key stakeholder is the BEMS teacher. Within this service the BEMS teacher developing the training had significant interest in the project, consequently the schools with which he worked were the first to be offered the opportunity to be involved. The author has worked in collaboration with this teacher to ensure that training and use of the behaviour plan reflect the intentions and purposes of the intervention. While the BEMS teacher was involved in the training and intervention development, he was not involved in analysis or collection of any results.

4.3 Philosophical Standpoints of Psychological and Educational Research

It has been suggested that, as research has developed over time, the number of paradigms and methods have increased, and are referred to with varying importance assigned to them throughout research literature (Mackenzie & Knipe, 2006). These research paradigms ("philosophical assumptions that guide and direct thinking and action" influencing the way an individual interprets the world (Mertens, 2005, p7)) are guided and defined by the philosophical, epistemological, and ontological assumptions and beliefs of the researcher. It is therefore important when conducting research to acknowledge biases and influences that the researcher's own belief system and guiding philosophies may have on the research. To do this there must be an understanding of the ontological and epistemological understandings influencing the research.

Ontology refers to the nature of reality (Mertens, 2005) focusing on the nature of the social phenomena being studied (Cohen, Manion & Morrison, 2011). Ontological assumptions impact on the choice and interpretation of epistemological presumptions which then influence methodological decisions, methods of data collection, and analysis (Mertens, 2005).

Epistemology refers to the theory of knowledge, the validity and certainty of that knowledge (Colman, 2006), and the way in which the knowledge is gathered and communicated (Cohen et al, 2011) or the "theory of how things can be known" (Robson, 2011, p525). Commonly recorded understandings and assumptions of key paradigms will now be discussed in more detail in terms of their ontology, epistemology, methodology, leading to a discussion of the standpoint of the current research.

4.3.1 Constructivism

The constructivist paradigm is often referred to as naturalist. Constructivism states that reality is socially constructed and fluid; therefore different for each individual (Cohen et al, 2011; Mertens, 2005). There is a complete rejection of the concept of an objective observer, and unlike other paradigms (see 4.3.2, 4.3.3) objectivity is not strived for

(Cohen et al, 2011). The constructivist paradigm does not believe in an objective reality as it is believed to be construed by each individual through social interactions (Robson, 2011). Methods tend to be qualitative, focusing on analysis of language and meaning. The focus of research within this area is to develop a deeper understanding of individual social constructions within specific settings and activities which can be compared, without the attempt to construct an objective conclusion around overall "truths" (Cohen et al, 2011; Robson, 2011).

4.3.2 Positivism

Positivism supports the existence of one "objective reality" (Mertens, 2005) and this paradigm indicates the world is "knowable" (Cohen et al, 2011). Researchers are seen as value free, neutral, and objective observers (Cohen et al, 2011) whose purpose is to search for causal explanations (Mertens, 2005). Given this belief in objectivity, the analyses and methods used for positivist research are based on those of the natural sciences. They tend to be quantitative and the context of the research is not relevant to analysis of results (Cohen et al, 2011).

This epistemology determines that genuine knowledge can only be advanced through scientific method providing the clearest possible view. Procedures must be clear and able to be repeated by other researchers to obtain the same results (Cohen et al, 2011).

4.3.3 Post Positivism

Post positivism often uses methods from natural sciences in a similar way to positivist science, but aims to strike a balance between the social and rational (Robson, 2011). A key epistemological difference is the acknowledgement that the unique subjective theories and knowledge of the researcher will influence observational data collection and analysis (Cohen et al, 2011; Mertens, 2005). There is an acceptance of the fallibility of the human researcher, therefore also fallibility of the evidence collected. Post positivism aims for objectivity while acknowledging absolute objectivity is impossible (Robson, 2011). Post positivism seeks to uncover a "truth" but accepts that due to human limitations science can only determine more or less likely explanations for phenomena (Mertens, 2005). The paradigm seeks objective truth through the use of

multiple studies to gather increasing amounts of evidence for theories to increase confidence in particular explanations.

4.3.4 Pragmatism

Pragmatism is not committed to any individual paradigm or philosophy but is instead guided by practical experience (Robson, 2011; MacKenzie & Knipe, 2006). The pragmatic approach recognises and acknowledges social, psychological, and natural assumptions, believing that truth and knowledge are changeable over time (Robson, 2011). Pragmatism acknowledges that all research contains flaws and research using a variety of methods and theories can build an evidence base in conjunction with one another (Robson, 2011).

Through use of a pragmatic approach the methods are guided by the research questions rather than individual researcher commitment to philosophies (Bryman, 2006). The pragmatic paradigm is reported to allow for opportunities to consider different assumptions, data collection, and analysis to best fit the "real-world" situations often in the form of mixed methods research (Mackenzie & Knipe, 2006). As discussed in section 4.5, the pragmatic paradigm guides this research, and implications of this for the present study will be further discussed (see 4.4, 4.5, and 6).

4.4 Design

A number of paradigms are often associated with different methods and research designs (see 4.3). Research within the constructivist paradigm is often associated with (although not strictly restricted to) qualitative methods, while research within the positivist paradigms is usually associated with (although not strictly restricted to) quantitative designs. Due to the extensive number of designs existing in the research literature, this chapter will focus on the designs used within the current research, having been selected based on the research questions (as guided by pragmatism (4.3.4)).

4.4.1 Fixed and Flexible Designs

Fixed designs are described by Robson (2011) as designs that have tight, pre-specified, unchangeable procedures prior to data collection. Fixed designs primarily rely on quantitative data (use of numerical data and statistical analyses). In contrast, flexible designs evolve during data collection (Robson, 2011) and typically use qualitative data (represented in words, picture, or icons and analysed through strategies such as thematic analysis). The fixed design is more closely aligned with the positivist paradigm (see 4.3.2) where the researcher aims to control variables in an attempt to strive for objectivity through pre-specified methods and data collection procedures. This is designed to minimise the researcher's own world view / interpretations from influencing the processes after the investigation has begun. Contrastingly a flexible design would be more likely to fit within paradigms that encourage more exploration and subjectivity, such as constructivism (see 4.3.1). These paradigms encourage the researcher to develop and alter strategies according to the participants and processes throughout the investigatory / exploratory process. Weaknesses and strengths have often been debated in detail for both these approaches and align with criticisms between paradigms.

A mixed methods approach is where a combination of fixed and flexible designs are used (Mackenzie & Knipe, 2006). Some researchers claim that a mixed method approach requires a "greater level of skill" leading to research that has a greater level of impact and understanding (Gorard & Taylor, 2004). The mixed methods approach is not restricted by paradigm and therefore does not risk wasting potentially useful data and information as it is disregarded based on the predominant paradigm or design adhered to

(Mackenzie & Knipe, 2006).

4.4.2 Randomised Control Trials

The randomised control design (RCT) is often described as the "gold standard" of experimental design (Robson, 2011; Cohen et al, 2011; Mertens, 2005). The RCT is used to establish causation; isolating and controlling independent variables to measure their effect on dependent variables. This design is described as a "true" experiment and suits laboratory experiments where researchers can exert high levels of control over environmental and experimental variables (Cohen et al, 2011). Groups are assigned randomly to ensure a greater likelihood of equivalence between control and experimental groups. The most common designs are pre- post, or repeated measures, however there are other variations of the RCT (Mertens, 2005).

The RCT has a history of being seen as the purest experimental design due to its ability (in laboratories) to establish clear cause and effect conclusions by eliminating confounding variables. Other strengths include not being susceptible to history, maturation, instrumentation, and mortality factors (Mertens, 2005) (see 4.6 for explanations of these terms of validity).

The RCT has been criticised for simplifying complex situations and is therefore often seen as inappropriate in educational settings (Pawson & Tilley, 1997). This view stipulates that studies investigating education must take place in educational settings, consequently being at risk of a number of confounding variables (Cohen et al, 2011). Contrastingly some researchers feel that these weaknesses can be addressed by analysing data of "sub-groups" to answer questions about "what works for whom in what contexts" (Pawson & Tilley, 1997). RCT's are at risk of a number of threats to validity such as the Hawthorne effect (see Table 4-1), Type I and Type II errors (see Table 4-1), and generalisability (see 4.6 for a more detailed discussion of validity and reliability threats).

4.4.3 Single Case Experimental Designs

Single case experimental designs (SCEDs) are the detailed investigation of individual cases where variables are manipulated and measured through quantitative measures. The SCED originates from behavioural psychology and compares a series of dependent variable measures during the baseline (no variable manipulation) of a single case to an intervention phase (an "intervention" or manipulation of a variable is introduced) (Cohen et al, 2011). The SCED often relies on observational measures as measures must be repeatable. This risks observer bias, especially where dependent variables are not adequately identified and defined (Cooper et al, 2007).

A SCED should allow for methodological replication (Horner et al, 2005). This allows each SCED to be compared to others within the field to contribute to understanding of groups, while focusing on subject variability (Barlow, Nock & Hersen, 2009). SCEDs have been shown to be particularly useful research tools for the SEN population due to the non-homogenous nature of the sample (Horner et al, 2005). As challenging behaviour can be so varied and the population of young people exhibiting challenging behaviours is not homogenous, SCED can contribute to understanding of challenging behaviour accounting for individual variability. Case studies can also be used to explore individual characteristics, often in greater detail than a SCED typically would (Robson, 2011). However, a case study has less focus on exploration of outcome data (Cooper et al, 2007) and is therefore less suited to the research questions posed by the current research.

SCEDs are at risk of being impacted by history effects due to the lack of a control comparison group. As SCEDs focus on non-homogenous samples, emphasising the individual (Horner et al, 2005), rendering direct comparisons to groups impossible. Consequently, the SCED compares two conditions for the same case and where possible aims for a "stable" baseline to increase reliability of analysis. The simplest design is the AB design (baseline, intervention), validity and reliability can be increased through designs that use withdrawal phases or multiple baselines (Cohen et al, 2011), independent and dependent variables will often dictate which designs are possible.

4.5 Epistemology and Design of the current research

The epistemology and design of the current research was developed based on the research questions, which were created based on needs of the local area and the intervention. Due to the impact of these factors on the planning of the research and the approaches taken within the research process, the epistemological approach adopted within this research is a pragmatic approach (Mertens, 2005).

The nature of the research questions requires 2 different designs with the designs best suited to their stage of inquiry. The pragmatist approach allows for "transgressing philosophical principles" (Bryman, 2006), allowing the researcher to use mixed methods. The designs were determined by the questions, rather than any paradigm preferences of the researcher (Bryman, 2006). However, it may be noted that researchers with a constructivist understanding (see 4.3.1) would suggest that the researcher preference for pragmatism is due to subjective preferences and philosophical constructs, therefore guiding the questions to lend themselves to particular designs. Although the researcher's theoretical and philosophical orientation may affect the interpretation of what best suits the research questions, the pragmatist approach appears to allow for more flexibility, considering all data gathered, even where collection methods or interpretive methods of a specific epistemology are not met (Mackenzie & Knipe, 2006).

Research question 1 is a group question focusing on teacher behaviour, based on a group independent variable (whole school training). As the questions are group based with specific independent variables a RCT design has been developed. This will inevitably lend itself to the quantitative and positivist / post-positivist paradigms (4.3.2 / 4.3.3). Data collection procedures are quantitative, however the researcher will also take note of any qualitative feedback given by participants. This acknowledges the constructivist epistemological standpoint, aiming to ascertain subjective views, individual interpretations, and perceptions. The pragmatist paradigm allows for this flexible data collection, which strict adherence to other paradigms would not allow for.

The second research question focuses on a largely heterogeneous sample (young people exhibiting challenging behaviours in school). This therefore suggests that a SCED design may be appropriate, allowing for an understanding of the independent variable on individuals with different characteristics / environments etc. (Horner et al, 2005). While a case study may have allowed for deeper investigation of issues surrounding individual participant behaviour (Robson, 2011) it would not have had a clear focus on exploring outcome data (Cooper et al, 2007). As the research question considers the impact of an independent variable on numerically obtainable dependent variables, the outcome focused SCED design seems appropriate (Cooper et al, 2007). Staff and young people will be in contact with the researcher which will allow the researcher to note any additional qualitative feedback given, allowing for some of the added depth a case study would provide, if appropriate.

In summary, this research aims to target research question 1 (a, b, c) through a RCT (with groups allocated as randomly as possible with regard to training slots available for each individual school). The research will target research question 2 through use of SCEDs.

For both research questions the data collection procedures are largely quantitative. Due to the pragmatist nature of research there will also be reference, where appropriate, to any additional ad hoc information given by participants providing it adds breadth and understanding to the analysis. This allows for attempts to provide objective scientific data and analysis while also acknowledging the social constructions of participants (and therefore to an extent the researcher).

4.6 Validity and Reliability

4.6.1 Validity

Validity in terms of scientific design refers to the amount a design can be relied upon to lead to replicable, controllable, practical, and objective conclusions (Cohen et al, 2011). For a method / instrument to be valid it must accurately measure that which it aims to measure (Field, 2013). Validity can be separated into internal and external validity. Internal validity refers to the quality of conclusions drawn from the data and how "true" they are within the population and methods used (Colman, 2006; Cohen et al, 2011). In comparison, external validity refers to the extent the research generalises to other populations / settings / research methods (Cohen et al, 2011; Colman, 2006). As all research is subject to threats to validity (Cohen et al, 2011) it is important for researchers to be aware of various threats so that design characteristics minimise these threats as much as possible.

Table 4-1: A table briefly summarising common threats to validity.

Threat	**Definition**
Internal Validity	The confidence with which relationships between variables can be reported.
History	Differences found may be due to differences in the conditions unrelated to the independent variable.
Maturation	Subjects may change between pre- and post-measures in ways unrelated to the independent variable.
Statistical Regression	Participants may regress towards the mean.
Testing	Subjects may be impacted by the testing itself.
Instrumentation	Tests or measuring instruments may be unreliable.
Selection Bias	Groups in different conditions may differ to begin with due to selection processes (may also interact with other threats to validity such as history and maturation).
Diffusion of treatments	Participants in a control group may be informed of the information / treatments shared with the experimental group therefore altering control group behaviour.
Experimental Mortality	Loss of participants due to drop out.
Type I error	Failure to find an effect where it exists.
Type II error	Finding an effect where no effect exists.
Construct Validity	Measures are measuring the construct that they are claiming to measure.
External / Ecological Validity	Generalisability of the results outside of the specific participants / settings involved.
Hawthorne Effect	Effects are due to participation in research rather than the independent variable itself.
Sample Representation	Poor sampling meaning the target population are not reflected in the research.

Table adapted from information in Robson (2011) and Cohen et al (2011).

Considering all threats to validity in great detail would be beyond the scope of this thesis. However, key threats to validity for the current research will now be discussed in some detail, with reference to attempts made to minimise these threats. The list below is not an exhaustive list of threats that may impact on the current research, but describes a few ways in which a number of the key threats are considered.

4.6.1.1 History and Maturation

This is a potential threat both to the validity of the SCED and the RCT. As the research takes place over a number of weeks there is a risk that the participants will be affected by their own idiosyncratic environments. It is unlikely that every school will experience history effects; by increasing the number of schools and participants the threat is minimised. Use of RCT aims to account for any history or maturation effects that may occur within / between schools. The author is unaware of any systemic changes within individual schools that may have impacted the staff during the 2 half-terms when the research took place.

The SCED participants may be impacted by history and maturity (Barlow & Hersen, 1984) during the intervention. This could mean that change detected in the dependent variables would be due to natural changes in the environment / individual over time, rather than due to the independent variable itself (Kratochwill et al, 2010). This could then risk leading to a Type II error. However, it was ensured that the participants would not be receiving any additional interventions / having their education altered during the intervention, compared to baseline. Consequently, the participants' own baselines serve as their control although trends must be interpreted with caution.

4.6.1.2 Testing / Instrumentation

The RCT is at risk of being impacted by effects of the instruments and re-testing procedures used with staff. Staff were required to complete the same questionnaires 2 or 3 times. This may result in staff answering differently the second / third time they completed the forms (possibly risking regression to the mean). Staff were more likely to have become aware of the design and specific purposes of the research by the second or

third completion of the questionnaire, possibly impacting responses.

The SCED instrumentation is at risk of observer bias or error (see Table 4-5) as the dependent variables are measured by staff (see 4.6.2 for further details). There is a possibility that use of these observation measures impact on the participant's behaviour, resulting in measuring the impact of the presence of the observation measures rather than the independent variable itself. Baseline measures are taken to attempt to minimise this impact and SCEDs would ideally have stable baselines before moving to the intervention phase.

4.6.1.3 External Validity

Threats to external validity are particularly prevalent in SCED designs. Yet, when interpreting SCED research it is important to consider the design purpose. The design does not claim to be generalisable, but focuses on the individuality of participants (Horner et al, 2005). However, SCEDs can be used to support and contribute to wider research in the field, therefore contributing to overall generalisability of the evidence base. It is crucial in SCEDs to ensure dependent variables, independent variables, and individual characteristics are accurately and clearly described so that they can begin to create descriptions of "who and under what conditions" (Horner et al, 2005).

The RCT uses a method that is more appropriate for generalisation (Robson, 2011), however caution is still necessary when generalising results as all schools are from the same county in the UK and recorded over the same time period. In order to increase generalisability schools with differing demographics were selected (see 4.7).

Participants of the SCED and RCT may be impacted by the Hawthorne effect. The SCED participants were being closely observed by teachers. Although students were not explicitly informed of the behaviours being recorded, or observational data collection procedures, as they were informed of the research it is likely they were aware of staff taking data. This may in itself have impacted data, although the baseline measures act as an accustomisation period for participants. Validity is strengthened when a SCED has a

stable baseline prior to intervention (however, see 4.11 for a review of issues regarding baselines within the current research). Despite the control group of the RCT not receiving direct input, they were made aware of being researched through completion of the questionnaires. This awareness of being a subject of research may impact their behaviour (Cohen et al, 2011) potentially resulting in their behaviour during the research period not reflecting standard unstudied attitudes and behaviours.

*4.6.2 **Reliability***

Reliability refers to the interpretability of research / instruments / data across settings and populations (Field, 2013). Reliable research designs and conclusions are "trustworthy or dependable" (Colman, 2006).

Table 4-2: A table briefly summarising common threats to reliability.

Threat	**Definition**
Stability	Repetition of the research would lead to the same results.
Internal Consistency	Instrumentation and data should be controllable, predictable, and replicable.
Observer effects	Risk of the effects being due to the presence of the particular observer rather than the independent variable.
Researcher objectivity	Interactions between the researcher and the participants may bias results therefore decrease objectivity and reliability.
Representativity of sample	The participants should be representative of the population the conclusions make assumptions about.
Observer Bias / Error	The observer may bias (or make errors in) the ratings taken, to either support or reject hypotheses.
Participant Bias	The participant may intentionally or unintentionally bias results based on own hypotheses / opinions.
Failure to accurately describe independent variables	This would render future replications impossible.

Table summarising information from Robson (2011) and Cohen et al, 2011).

Below are summaries of how a few key reliability threats relate specifically to the current research. Although the researcher must be aware of and ensure controls are put in place in regards to all threats to reliability, those of particular concern to the current research methodology are highlighted.

4.6.2.1 Observer Effects and Failure to Accurately Describe Variables

Observer effects may impact the data of the SCED as the student behaviour may be impacted by being observed. However, in the A-B SCED design the repeated measures in the baseline period are designed to decrease observer effects on participants, as they become accustomed to the data collection procedures. To minimise effects of observer on the students, class staff collected the data, this resulted in the number of staff in the room not needing to alter for data collection. However, student behaviour may have been impacted by becoming aware of staff observing them more closely. Variables need to be accurately described so that behaviours can be accurately defined, identified, and recorded. In order to ensure data fidelity, data checks were taken in each classroom by the researcher and 100% agreement was found in each case, suggesting that variables were accurately defined and identifiable. However, errors may still have occurred due to the nature of a busy classroom. The students appeared to be significantly affected by researcher presence in the classroom with almost 0% exhibition of behaviours on each check, due to this the number of checks was decreased.

4.7 Recruitment and Allocation of Participants

4.7.1 Question 1

The 32 primary schools in the area that worked with the BEMS were offered the opportunity to be part of the research. An initial e-mail was sent explaining the purpose of the research and detailing requirements of involvement. The school EPs and the BEMS teacher also spoke to schools, on the researcher's behalf, after the initial e-mail. A week later a further e-mail was sent to serve as a reminder of the request for school involvement. From these initial e-mails and discussions 11 schools expressed interest in taking part in the study (1 later dropped out due to organisational changes). To increase the number of schools involved in the research, the researcher also offered the opportunity to appropriate schools within the patch of schools she worked in. Of these schools 2 agreed to take part in the research and an additional school took part in the pilot study.

Of the schools taking part in the final research design 10 schools were from the rural sample and 2 from the urban sample. The sample was taken from a Northern England county with a mixed demographical population, including rural and urban coastal areas.

The 10 rural schools (9 primary, 1 infant) had between 3 and 8 members of teaching staff. The number of pupils in the control group schools ranged from 24 to 248 (8-12% free school meals eligibility). The number of pupils in the schools in the experimental group ranged from 12 to 63 (0-44% free school meals eligibility). The urban schools were from a deprived coastal town. The urban school in the experimental group has approximately 45 members of staff and 320 pupils (56% free school meals eligibility), the urban control group school has approximately 35 members of staff and 450 pupils (40% free school meal eligibility).

4.7.2 Question 2

A school from the rural and 1 school from the urban areas agreed to take part in the SCED. The rural school has approximately 45 pupils (45% eligibility for free school meals), and the urban school has 320 pupils (56% eligible for free school meals). These schools were sent initial information by email and the researcher met with the head teacher / lead behaviour teacher for each school to discuss potential participants. The schools were asked to identify 3-4 pupils who met the inclusion criteria (Table 4-3) and then obtain pupil and parental consent. Both schools identified 3 participants, however parental consent was not granted for 1 participant in the urban school, resulting in 5 participants. The researcher met with the teachers to discuss specific behaviours and data collection methods (see below for participant descriptions). Target measurable behaviours were discussed and data collection methods for each were devised including definitions (see SCED measures for more details 8.8.2).

Table 4-3: A table showing the criteria for student participants for the research.

Criteria	Definition
Challenging Behaviour.	Behaviour that challenges the individual or the staff / pupils around the individual. For the purpose of this research the behaviour had to be observable and measurable.
Ability to discuss the behaviour.	The participant was required to have enough verbal ability and understanding to be able to discuss their behaviour.
Be able to attend weekly meetings.	The pupil was required to attend weekly meetings with the staff member.
Same educational provision planned for baseline as in the intervention.	The educational provision was required to be consistent with the same level of support from outside agencies and within school interventions throughout baseline and intervention phases.

The following cases were selected:

S1a (Rural school); S1a is a Year 5 male on the SEN register in school. He is in a class of 12 Year 3, 4 and 5 pupils. S1a receives a high level of writing intervention in school. S1a has a history of attention seeking behaviour that disrupts the class. These behaviours most commonly included putting his hand up, making loud displeased sounds if he was not chosen, or not knowing the answer when he was chosen. S1a's teacher reported that S1a often refused to complete his work unless he was receiving 1:1 support. The behaviours recorded for S1a were inappropriate reactions to not being chosen, interrupting the teacher, complaining about lack of support with familiar work and refusal to complete work (see Appendix 7 for full behaviour definitions).

S1b (Rural school); S1b is a Y3 male in the same class as S1a. S1b is on the schools SEN register and reportedly struggles in all academic subjects across the curriculum compared to same age peers. S1b has a history of regular medical appointments although these have not led to any diagnoses. School staff report that S1b tended to be very quick to cry at what seemed to be very small incidents. This crying was often disruptive for his class as well as getting in the way of his own success (socially and academically). The behaviours recorded for S1b were inappropriate crying, and inappropriate noises during work time (see Appendix 7 for full behaviour definitions).

S1c (Rural school); S1c is a Year 2 male in a class of 14 Year 1, 2, and 3 students. He joined the school in September 2013 after being excluded from his previous school due to acts of physical aggression and repeated theft. He is on the school action plus register and receives play therapy once a week (provided by the BEMS). At home S1c has been a history of aggressive behaviours and the family were receiving support from a number of agencies as part of a "team around the child" throughout the research. S1c's behaviour in school was reported to be uncooperative, avoiding tasks and demands. The behaviours being recorded for S1c were the number of times he refused to carry out instructions (and length of time of each occurrence).

The data for the two participants from the urban was not included in the analysis as their data did not meet the inclusion criteria (Table 4-3):

S2a (Urban School): S2a is a Y2 female in a class of 30 students. S2a has a history of hitting other students when she is angry. S2a's teacher reports that these occurrences often led to S2a refusing to accept responsibility and to stop engaging with staff/peers for extended periods of time. S2a is reportedly intelligent and is achieving levels above the typical norms for all academic subjects. The behaviour being recorded for S2a was "hitting". The data for S2a was not included in the analysis as there was not enough data (irregular behaviour) to provide meaningful analysis after the intervention stage.

S2b (Urban School); S2b is a Y4 male in a class of 32 of students. S2b is on the SEN register and achieving below average across the curriculum. S2b has difficulties with his peers. S2b was reported to often shout out questions and comments in class during taught input. The behaviours being recorded for S2b were shouting out in class and crying. Due to illness S2b was not in school for the majority of the time during intervention (only one whole week of the half term). Consequently the data for S2b has been removed from the analysis (see Table 4-3).

4.8 Pilot Study

A pilot study was carried out to gather feedback on, and trial timings of, the training session. A phone discussion was had with the special educational needs coordinator (SENCO) from the school to arrange a date for the training. Staff attended training on a voluntary basis.

4.8.1 Participants

The participant school for the pilot study is a Roman Catholic primary school within the urban area where the research took place. The school was offered this opportunity as the school receives allocated EP time from the researcher. The SENCO offered the training to all teaching staff in the school; 17 staff agreed to take part (3 PGCE students, 5 teaching assistants, 6 teachers, SENCO, deputy head teacher, and head teacher).

4.8.2 Design

The pilot study consisted of evaluation of the training package developed by the researcher (see final training package in Appendix 6). The training was delivered to volunteer staff members and staff were asked to complete an evaluation form (Appendix 3). The pilot school was only able to provide 1 hour for the training session.

4.8.3 Results of Pilot Study

The pilot study showed that the training fitted comfortably within the 1 hour slot. The evaluation forms were analysed and key points were highlighted, changes were made accordingly (see Appendix 4 for specific evaluative feedback). The changes made as a result of the feedback included adding examples and making the behavioural functions and technical terms sections clearer. Some staff felt that more time or a slower presentation would be beneficial. Due to the times available in the schools an extra half hour could be added for the research project, some of which would need to include considering example scenarios, responses, and behaviours, (as suggested by 11 staff). 2 staff members felt that more interactive activities would benefit the training and these were included where time was available in training sessions (e.g. practising sections of the behaviour plan with one another). Overall, the feedback was positive and staff

reported finding the information and the behaviour plan useful. This suggests that the training was suitable for its purposes of supporting and educating school staff about behavioural theory and a related intervention.

4.9 Intervention

The intervention was developed by the BEMS specialist teacher due to his belief that a number of young people with challenging behaviour referred to his service would benefit from a solution focused and pupil led intervention. The intervention had been put in place by the BEMS teacher with a few individual pupils. However, due to the regular visits required (at least weekly) to speak with the pupil it was felt it may be more appropriate for school staff to lead the intervention. Consequently a training package was developed by the researcher in collaboration with the BEMS teacher with the aim of equipping teachers with the knowledge, skills, and materials to implement the intervention. Some amendments were made to the intervention in light of the planned research (e.g. example questions were added to the behaviour plan to provide guidance for the staff members).

The intervention's development is based on a number of key behavioural psychology, solution focused and person-centred planning principles. A weekly discussion between pupil and staff member is had and recorded on sheet 1 of the behaviour plan (Appendix 5). The training intervention sheet includes example questions based on cognitive behavioural principles, and functional analysis to define challenging behaviours and the students' emotional and behavioural responses, while identifying potential antecedents. This then leads to a discussion around alternative behaviours that the student may find helpful to replace the challenging behaviour from the previous section (based on behavioural principles). This section includes solution focused questions that aim to encourage the student's motivation and belief that they can act differently in given situations.

Following this weekly discussion, the staff member and pupil set a weekly target (developed as an outcome of the discussion). This target setting should be led by the pupil, although staff support may be required to ensure targets are realistic and achievable. The staff member and pupil then record the relevant behaviours (according to target). Pupil chosen rewards will be available if targets are met at the end of the week (the timescale for rewards is changeable according to the behaviour and the individual characteristics of the student) (see Appendix 5).

4.9.1 Developing the Training Package

The training session (Appendix 6) was devised by the author after meeting with the BEMS teacher to discuss the implementation of the intervention and the guiding psychological theories. Although the training content was developed by the author, the training was discussed with the BEMS teacher on a number of occasions and alterations were made accordingly.

Research into the training of school staff indicates that training should be directly linked to strategies that can be used in the classroom (this may explain why pilot participants felt the behavioural intervention part of the training was the most useful (Appendix 4)) (Bubb & Earley, 2013). Artman-Meeker & Hemmeter (2013) found that giving a rationale for a type of practice, with sample materials, and examples of its use, led to successful implementation of the practice in school. Research suggests that teachers benefit from global theoretical training as well as training in specific intervention implementation (Jennet et al, 2003). This training package therefore included a section on the theory behind the intervention to provide a clear rationale for its use. Although examples of practice and behaviours were used in the training, due to confidentiality previous examples of specific behaviour plans could not be given.

The pilot study led to the alteration of some elements of the training such as clarification of technical terms and the addition of example behaviours and optional interactive activities.

4.10 Measures

4.10.1 Question 1

Teacher burnout and teacher self-efficacy were measured through repeated questionnaires. The questionnaire measuring teacher self-efficacy in behaviour management was the Teacher Efficacy in Classroom Management and Discipline Scales (TECMDS) (Emmer & Hickman, 1991). The TECMDS has often been used in self-efficacy research and its development was based on the Teacher Self-Efficacy Scale (Gibson & Dembo, 1984), which also includes general instructional perceived self-efficacy (Brouwers & Tomic, 1999). This research uses a questionnaire based on self-efficacy specifically linked to classroom management (TECMDS), as it has been shown to be a different construct from general instructional self-efficacy (Emmer & Hickman, 1991). The TECMDS has been used throughout research with a number of education staff including pre-service teachers (Emmer & Hickman, 1991), qualified mainstream teachers (Brouwers & Tomic, 2001), and special education staff (Leyser, 2007). To the author's knowledge it's reliability has not been specifically tested for teaching assistants (TA's), potentially impacting appropriateness of using the measure with TA's. However, given the apparent lack of literature and call for self-efficacy scales relating to TA self-efficacy (Higgins & Gulliford, 2014), shared responsibilities for behaviour between TAs and teachers for behaviour management in schools, and the domain-specificity of the TECMDS, it was deemed the most appropriate measure for this research. As the research question also focuses on whole school outcomes it was felt it would not be appropriate to have different measures according to different roles.

Due to the reliability measure of the TECMDS being based only on the scales related to the independent variables of this research it seemed more appropriate than other self-efficacy scales that are popular within self-efficacy literature (e.g. Tschannen-Moran et al, 1998). Using scales that include additional self-efficacy scales would either result in using increased amount of staff time to complete potentially irrelevant questions, or potentially impact reliability of scales by using only sections rather than the whole scale on which reliability scores were based.

The TECMDS have been used in a number of research studies mentioned in the systematic review (see 1.4) and literature review (See 1) and have been found to have a 0.79 reliability score (Emmer & Hickman, 1991).

The questionnaire used to measure teacher burnout is the "Shirom-Melamed Burnout Measure" (SMBM) (Shirom-Melamed Burnout Measure, 2005). Although the literature included in the systematic review (see 1.4) and the literature review (See 1) suggest that the "Maslach Burnout Inventory" (MBI) (Maslach et al, 1996) is the most popular teacher burnout scale in the field, there are a number of issues with the use of this measure. MBI is American and the availability of the measure is very limited (with significant financial commitments). Although this measure has been the most popular measure within burnout literature (Shirom & Melamed, 2006), 3 other measures have also become popular in the evidence base (Qiao & Schaufeli, 2011). Qiao & Schaufeli (2011) compared the 4 most popular burnout measures and found they all had a high Cronbachs alpha for overall and individual burnout factors and scores (ranging between 0.67 and 0.94). All 4 measures were recorded to have high construct validity. Shirom & Melamed (2006) compared the MBI and the SMBM across different occupational groups and found that both had equivalent high rates of reliability and construct validity.

4.10.2 Question 2

The measures for question 2 are teacher (and class staff) observational data. SCEDs often require observational data due to their dependence on repeated measures (Cooper, Heron & Heward, 2007). Target behaviours were defined (Appendix 7) in collaboration with the class staff and examples and non-examples of behaviour were included to increase accuracy (Cooper et al, 2007). During the baseline and first 2 weeks of the intervention period the researcher also sat in the classroom and collected data for 10 minutes every 2 weeks (although this was ceased due to apparent impact of researcher on the students). Inter-rater reliability was 100% suggesting that the behaviours were accurately defined. Behaviour data was taken daily throughout the day for each student. (See Appendix 8 for example data collection sheet).

4.11 Ethics

The researcher submitted an ethics proposal to the ethical approval committee for the university. This was returned and necessary alterations made and re-approved prior to beginning the research (Appendix 9). Any small changes to the study were approved by the university supervisor who was continually made aware of these. Research in educational settings presents a number of ethical considerations given the potentially vulnerable sample (both young people and staff), potential impact on the environment for participants and others in the setting, and potential impact on learning. A number of ethical considerations and how they were managed will be discussed below.

Prior to data collection, consent forms were completed by staff taking part (Appendix 10 and Appendix 11), carers of the young people participating in the SCED (Appendix 12), and the young people (Appendix 13). To ensure participants were giving informed consent, the young people were guided through the consent form by a staff member. Students with questions were given the opportunity to discuss these with staff and the researcher before any data was collected. Phone calls took place between the researcher and any parents / carers who had questions. Staff were approached via the contact person for each school (SENCO or head teacher) and given the consent forms and questionnaires to complete. The experimental group were all given time to ask questions and the research was re-iterated by the researcher prior to the initial training (where the consent forms and questionnaires were collected). Staff members were at this point reminded that they were free to withdraw at any time and given the researcher's contact details so they could contact the researcher without senior staff being aware.

The control group did not receive this talk prior to collection of consent and questionnaires. However, on collection of the consent forms, despite the consent form including researcher contact details and explicitly stipulating that the researcher was available for questions, some staff reported they did not feel there was adequate opportunity to ask questions. As a result of this, no data was analysed until the researcher had visited the control schools (usually linked with collection of follow up questionnaires) to further explain the research, and re-emphasise the anonymity and right to withdraw.

All data has been stored securely by the researcher and reported anonymously. Only the researcher has access to information identifying staff members and participants. Identifying data was stored in case a member of staff / student wished to see their data, or if any safe-guarding / welfare concerns arose. Should these be required the participant would be contacted immediately and made aware. All questionnaire responses were stored using number codes, these were stored separately to names of participants. Neither school nor participant names are reported to ensure confidentiality and anonymity.

4.11.1 Question 1 Specific Considerations

Due to the staff questionnaires measuring personal constructs some ethical considerations were raised. The personal nature of the research emphasised the importance of informed consent. Self-efficacy and burnout measures given to staff may cause stress or bring personal issues to light after their own reflection. Staff were assured that they could contact the researcher at any time / withdraw from research / ask for support regarding this. When data was analysed any considerable outliers were to be contacted directly by the researcher and offered support, or assistance in obtaining support. As a result of this 1 staff member was contacted, but no further support was requested.

Staff may have felt pressured to attend whole school training for the research. However, the EPS for whom the researcher works often led training in schools and therefore similar training would be likely to take place in schools. This training could be considered as continuing professional development training which school staff are required to undergo every year. The training itself was likely to benefit staff knowledge and therefore unlikely to have any ethical implications of its own.

Staff were not explicitly told the research hypotheses until after data collection when staff were de-briefed. Staff were reminded that they could request their own personal results if they chose, after analysis (right to withdraw was reiterated).

4.11.2 Question 2 Specific Considerations

The participants for the SCED were aged between 5 and 11 years. This therefore automatically results in their being classified as vulnerable participants. The young people and their carers were asked for informed consent before the research (Appendix 12 and Appendix 13). It was ensured that participants understood they could withdraw their data / consent at any point without giving a reason. The students would have been receiving direct input from staff members with whom they are in daily contact regardless of this research. This should decrease any direct impact of being part of the research on them. Regular fidelity checks of the intervention were scheduled by the researcher, giving the researcher opportunity to ensure the intervention was not causing any harm.

Staff were given access to student data as part of the intervention. This data would usually be available to staff in schools. For the purpose of research, student data was coded with a confidential reference code. The data was stored anonymously, thus ensuring confidentiality. If a student disclosure were to occur, the usual safeguarding protocols within the school and county council would have been adhered to.

As this research focuses around managing challenging behaviour there was potential risk to staff and students. However, these staff would have been managing this behaviour within their everyday jobs despite the research. Implementation of any behaviour management strategies can risk increased behavioural outbursts. Therefore, staff were supported by the researcher throughout, and able to contact the researcher with any concerns. Staff were reassured that in case of increased risk to safety (above what is usually dealt with as part of a regular school experience for that student) where the intervention no longer appeared to be appropriate, the student would be withdrawn from the research. Individualised support was available throughout the research to the staff and student by the TEP or other appropriate professional services the school / TEP would normally be able to access regarding such matters.

Students' mood may have been impacted by the intervention. They may have found the

initial stages stressful as they were required to talk about their own behaviours. However, in a school setting where a student has behavioural difficulties these are often referred to / discussed. The intervention aims to protect students from negative discussion by requiring positive and non-judgemental discussions about behaviour.

A SCED design requires a "baseline phase" where data is taken without the intervention being in place. The minimum number of measures for a baseline phase are 3 (Barlow & Hersen, 1984), but the baseline is supposed to continue until it is stable. Although a baseline phase is required to inform those involved whether an intervention is effective (which can therefore avoid students being exposed to ineffective interventions for longer than necessary), waiting for a stable baseline may risk withholding potentially beneficial interventions. Consequently, although stable baselines were not achieved for the behaviours, the intervention was put in place after 2-3 weeks of baseline. Whether intervention was put in place 2 / 3 weeks after baseline depended on the school's own time schedules, and the impact of the behaviour on the staff, participant, and other students.

4.12 Procedure

4.12.1 RCT (Research Question 1)

Research question 1 (see 2) was researched through a RCT. Table 4-4 shows the independent and dependent variables for this question. 12 schools took part in the research (6 in each group) equalling 108 staff (60 in the control group and 48 in the experimental group). The number of questionnaire responses (see Appendix 14 for example questionnaire) gained from the experimental group were 60 at time 1, 25 at time 2, and 35 at time 3. The control group responses totalled 60 at time 1, and 37 at time 3.

Table 4-4: A table showing the independent and dependent variables for research question 1

Research Question		**Name of Variable**	**Levels**
Question 1a	Independent Variable	Time	Baseline, follow up six weeks later.
	Dependent Variable	Number of children the behaviour plan is used for.	
Question 1b	Independent Variable	Condition	Control Experimental
		Time	Baseline, Six weeks after training, follow up at six weeks.
	Dependent Variable	Self-efficacy Scores	
Question 1c	Independent Variable	Condition	Control Experimental
		Time	Baseline, Six weeks after training, follow up at six weeks.
	Dependent Variable	Teacher burnout Scores	

After participants had been recruited they were semi-randomly allocated into 2 groups (experimental and (waiting list) control) (see 4.7). Each school was allocated a code according to group. Training dates were agreed with the head teachers / SENCOs of the experimental schools, and schools were informed that they would need to complete the questionnaires immediately before training, approximately 1 week after training, and again 6 weeks later (see Figure 4-1 for timeline). Once these dates had been assigned, the control schools were contacted and dates for their training (after follow up data collection) were organised. The control schools were informed of the dates they would be required to complete their questionnaires (these were organised to ensure the number of schools completing questionnaires each week would be similar).

The training (see 4.9.1 and Appendix 6) was delivered to all consenting teaching staff in the experimental group. On completion of the training, staff were informed that they could contact the researcher at any time with any questions or for support relating to training. SENCOs / head teachers were emailed 6 (school) weeks later to share the next questionnaires and agree a collection date. Control groups were also contacted to complete questionnaires and arrange collection 6 weeks after the initial questionnaire. Control groups then received their training.

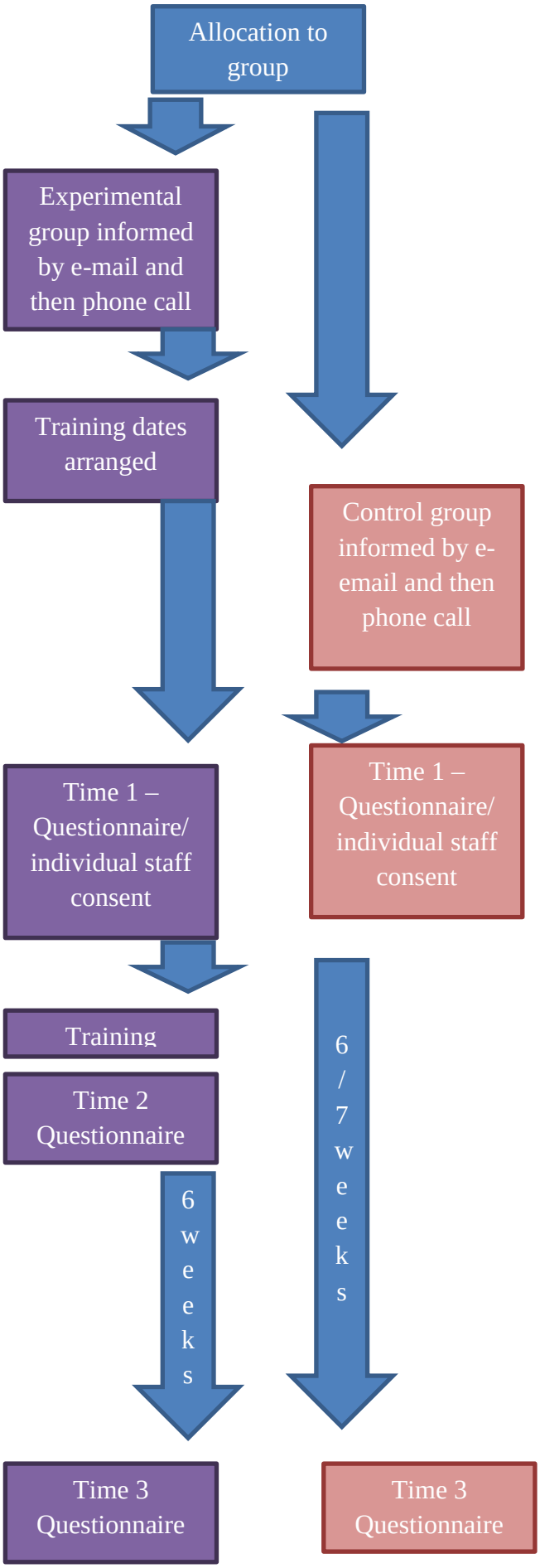


Figure 4-1: A diagram showing the timelines for research question 1 (start points were different for each school according to training date / date of completion of questionnaire 1)

4.12.2 SCED (Research Question 2)

Question 2 was researched through SCEDs (Table 4-5 shows the independent and dependent variables for this research question). 5 pupil participants from 2 (3 pupils from a rural and 2 from the urban school in the experimental group) schools took part in the research over 8 weeks. Due to a number of difficulties throughout the intervention only 3 participants (all from the rural school) are included in the analysis.

After recruitment of participants (see 4.7) meetings were held between the teachers / teaching assistants of the individual pupils and the researcher. In these meetings staff described the behaviours they felt were most disruptive to each individual pupil. It was decided which behaviours would be recorded daily (according to measurability, manageability, and influence of the behaviour in the pupil / other pupils). This included developing definitions (Appendix 7) and determining the methods of data collection. These details were then transferred onto a data collection sheet (Appendix 8) and the teachers began taking baseline data. Data collection fidelity checks began where the researcher gathered data in class every week, to compare data collected with the staff member. However, as researcher presence appeared to impact the student behaviour (for ethical reasons the students knew the researcher prior to data collection) these fidelity checks were stopped. The intervention was observed by the researcher every 2 weeks for each participant to ensure reliability of the intervention itself (Appendix 15). These checks showed that staff consistently followed the behaviour plan and incorporated each section of the behaviour plan into the discussion with the students using questions guided by the theories and example questions that had been provided in training.

Table 4-5: A table showing the independent and dependent variables for research question 2

Question 2	Independent variable	Use of Intervention	Pre and post Repeated measures
	Dependent Variable	Challenging behaviour frequency	

5 Results

This chapter will present the findings relevant to the research hypotheses. Findings are presented for each research question in turn. Each section begins with a discussion of pertinent aspects of analysis including rationale and description of the analysis procedures. Results for question 1 are presented in subsections according to the hypotheses. Results for research question 2 are presented case by case.

5.1 Research Question 1

1) In a sample of rural primary schools can training based on a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms:

a) Lead to use of the behaviour plan?

Hypotheses

a(i) Participants' receiving training will have increased their use of behaviour plans 6-8 weeks later, compared to participants not receiving any training.

a(ii) Participants' will have implemented the behaviour plan from the training session 6 – 8 weeks later.

b) Increase overall staff self-efficacy in behaviour management?

Hypotheses

b(i) Participants' self-efficacy scores will be increased after training and this will be maintained 6-8 weeks later.

b(ii) Participants' receiving training will have a greater increase in self-efficacy after training (6-8 weeks) than participants who have not received any training.

c) Decrease teacher burnout?

Hypotheses

c(i) Participants' burnout scores will be decreased after training and this will be maintained 6-8 weeks later.

c(ii) Participants' receiving training will have a greater decrease in burnout scores after training (6-8 weeks later) than participants who have not received any training.

5.1.1 Data and Analysis Plan

The data gathered for this question is all quantitative, and therefore will be analysed using statistical analysis designed for quantitative research. Participant attrition has resulted in the number of participant responses being different for each phase of the research. Before determining the type of statistical analysis that will be appropriate for this data (parametric or non-parametric tests), the data must be assessed for meeting assumptions of linear models (Field, 2013). This section will describe the necessary pre-tests and their purpose prior to identifying and executing the analysis procedure for each research question in turn.

5.1.1.1 Type of Data

The data gathered for research question 1 includes scale, interval, and nominal data. The dependent variables (burnout and self-efficacy) are measured through Likert scale questionnaires. These scales assign a number to statements such as “strongly agree” and “strongly disagree” using a numerical scale. There is an on-going debate in the literature about whether Likert data can be treated as scale data for the purpose of analysis or whether it must be treated as ordinal data (Purdey, 2013). It has been argued that Likert scale data is ordinal due to numerical allocation to steps within a subjective scale. This interpretation states that although the scale can be determined to increase, the interpretations of the difference between the numbers may not be equal within or between participants (Camparo, 2013). Specific analysis procedures have been developed for ordinal data (Camparo, 2013). Contrastingly, other researchers argue that the mathematical concepts of a Likert scale are important to the scale, and analysis of the data using methods appropriate for scale data is therefore appropriate for Likert scales. However, conclusions should be drawn with acknowledgement of these limitations (Purdey, 2013). In an article reviewing Likert scale analysis, Carifio & Perla (2007) state numerical and parametric methods analysis leads to clearer and more concise results than when analysed as if it is ordinal data. Due to this argument, assuming the data meets the assumptions required, this research will analyse the Likert scale scores as scale data. Descriptive data will also be presented to clarify meaning and provide visual representations of the data (Howitt & Cramer, 2008).

Demographic data has also been collected for these hypotheses in the form of nominal data (e.g. gender, yes / no questions), and categorical data (e.g. job role).

5.1.1.2 Scoring the Dependent Variable Scales

The self-efficacy and burnout scale scores were inputted into Microsoft Excel (2010). The self-efficacy scale included reversed items, these were altered to account for their reversal using Excel. The totals of the self-efficacy scores were then added together for each subscale ensuring Excel did not give totals for scales where data was missing by using "IF" statements within the equations. The subscales external (belief in external influences impacting on student behaviour), personal (personal beliefs in behaviour management abilities), and general (personal belief in general teaching abilities) were then added together to create the overall self-efficacy score.

The burnout scores were organised according to subscale (emotional exhaustion, cognitive weariness, and physical fatigue). The scores for burnout are calculated using the mean for each selection of scores within the subscales. The overall score is calculated using the mean for the 3 subscales. Again "IF" statements were used to calculate these to ensure total scores were only calculated where each item was answered.

Due to a number of schools having difficulties with printing questionnaires at phase 1 there were a large number of missing responses (especially for self-efficacy question 13) in phase 1. A number of other questions were also missed by a number of participants (see 5.1.2.2). Consequently 2 sets of scores were calculated for each participant questionnaire; an "actual" total as above (leading to a number of missing participants – especially in phase 1) and an "estimated" total. This used the average of the scores for the individual participant within the subscale to generate an estimated response for the missing item. Where participant number was significantly affected by individual missing item responses analysis was conducted for actual and estimated scores.

5.1.2 Demographics of the Sample

Prior to any data analysis the participant demographics will be presented. Missing variables will also be highlighted to ensure transparency of data analysis procedures.

5.1.2.1 Participants for Each Stage

Participant attrition occurred for both phase 2 and 3 of the data collection procedure (see 6.5 for discussion of this issue).

Table 5-1 shows the number of teaching staff for each phase, including group demographics. Where a staff member wrote multiple job titles the title considered most specialist or most senior was used, as all senior or specialist staff included were also qualified class teachers. The job titles have been categorised for clarity; Senior Leadership Team (SLT) (head teachers, deputy head teachers, SENCOs, and specialist lead behaviour teachers); Qualified Teachers (class and supply teachers); and Teaching Assistants (TAs) (advanced (ATA) and general teaching assistants (GTAs)).

Table 5-1: A table showing the number of participants in each phase

	Time 1		**Time 2**	**Time 3**	
	Experimental	*Control*	*Experimental*	*Experimental*	*Control*
Current Role					
SLT	7	9	7	7	6
Qualified Teacher	22	26	12	19	20
Teaching Assistant	19	25	6	9	11
Gender					
Male	5	9	4	5	6
Female	43	51	21	30	31
Age (years)					
20-29	9	15	5	7	11
30-39	15	16	9	12	11
40-49	9	13	4	7	6
Over 50	15	16	7	9	9
Number of years teaching experience					
0-1	3	4	3	3	2
2-5	15	18	7	11	14
6-10	7	8	5	5	1
11-15	8	10	1	6	5
16-20	4	10	2	2	8
21-25	4	6	2	2	5
Over 25	6	3	5	6	2
Unknown / Missing	1	1			
TOTAL	**48**	**60**	**25**	**35**	**37**
	108		**25**	**72**	

5.1.2.2 Missing Item Responses

There was a printing error for the self-efficacy questionnaires for a number of schools at time 1. This resulted in a reduced number of participants responding to question 13 (49 participants), thus resulting in no "general self-efficacy" score for these participants. Table 5-2 summarises any items missed by 5% or more of the participants. Where the missing data significantly impacted on the number of participant responses available for analysis, a second analysis was also conducted using estimated scores (see 5.1.1.2).

Table 5-2: A table to show any missing responses by 5% or more of respondents for each phase of data collection.

	Time 1 Number of participants (%)	**Time 2** Number of participants (%)	**Time 3** Number of participants (%)
Self-efficacy			
Q8 If one of my students couldn't do an assignment I would be able to accurately assess whether it was the correct level of difficulty.		3 (11.1%)	
Q11 There are very few students that I don't know how to handle.	9 (8.3%)		
Q13 When a student is having trouble with an assignment, I am usually able to adjust to his/her level.	37 (34.3%)	3 (11.1%)	

5.1.3 Parametric Assumptions

Most statistical tests are based on linear models (Field, 2013). In order for the data to be suited to parametric tests it must meet certain assumptions to minimise bias. If data does not meet parametric assumptions, non-parametric alternatives must be used. The parametric assumptions required are normality of distribution, homogeneity of variance, and equality of groups. Certain procedures also require additional assumptions to be met such as sphericity, which will be reported when relevant to each test. This section will briefly summarise these assumptions.

5.1.3.1 Assumption of Normality

Parametric tests of significance often assume that data are normally distributed. The Shapiro-Wilks test of normality was used to analyse data for assumption of normality (to determine whether data is appropriate for parametric testing). Any scores showing a significant value ($p<0.05$) indicate deviation from normality. See 5.1.4.1 and 5.1.8.1 for results of the tests of normality for the hypotheses.

5.1.3.2 Homogeneity of Variance

Homogeneity of variance determines whether the variance is similar for the groups. The Levene's test of variance can be used to assess the variance in data samples. Results that are significant ($p<0.05$) suggest a significant difference in variance between groups. See 5.1.8 and 5.1.9 for results of these tests.

5.1.3.3 Effect Size

Effect size measures the magnitude of the observed effect using standardised variables (Field, 2013). The magnitude of effect size is categorised according to the value of the partial eta squared. A partial eta squared of 0.01 indicates a small effect size, 0.06 indicates a moderate effect size, and 0.14 indicates a large effect size (Pallant, 2013).

5.1.4 Research Question 1a – Analysis and Results

1) In a sample of rural primary schools can training based on a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms:

a) Lead to use of the behaviour plan?

Hypotheses

a(i) Participants' receiving training will have increased their use of behaviour plans 6-8 weeks later, compared to participants not receiving any training.

a(ii) Participants' will have implemented the behaviour plan from the training session 6 – 8 weeks later.

Table 5-3: A table showing the independent and dependent variables for research question 1a.

Hypothesis	**Independent Variable (levels)**	**Dependent Variable**
Participants receiving training will have increased their use of behaviour plans 6-8 weeks later, compared to participants not receiving any training.	Group (2) Time (2)	Number of behaviour plans implemented by the participant.
Participants will have implemented the behaviour plan from the training session 6 – 8 weeks later.		Implementation of the behaviour plan. Intention to implement the behaviour plan.

For analysis of this research question the independent variables and dependent variables are shown in Table 5-3. Due to the specific nature of 2 behaviour teachers' classes (students are removed from the mainstream class and go to the behaviour class) 100% of students in their classes had behaviour plans pre- and post-training, thus rendering their data meaningless in relation to impact on number of behaviour plans in place; their data was removed for the purpose of this analysis. A further outlier was removed from

the control group, the teacher had responded “20” (all students) as the number of students with a behaviour plan. The wording may not have clarified to the teacher that by “structured behaviour plan” the research concerns itself with “additional” behaviour plans rather than the structured whole class behaviour plans, although a few staff referred to these in their responses (usually as addendums).

5.1.4.1 Assumption of Normality (Q1a)

The dependent variable “number of behaviour plans in place” was tested for normality of distribution. Tests were carried out for both groups and each time independently. All of the groups did not meet the assumption of normality according to the Kolmogorov-Smirnov test of normality (see Table 5-4). Visual analysis of the Q-Q plots suggest a distribution close to normal distribution (Appendix 16).

Although data is not normally distributed it is reported that ANOVA tests (test most suitable for analyses with the characteristics of the independent and dependent variables within this hypothesis) are robust to violations of assumptions, meaning they are still more reliable than non-parametric tests in many cases (Pallant, 2013). Consequently further pre-analysis checks will be conducted to determine if an ANOVA is still possible.

Table 5-4: A table showing the results of the Kolmogorov-Smirnov tests of normality.

		Statistic	**Df**	**Sig**
Time 1	Experimental	0.27	22	0.00
	Control	0.32	26	0.00
Time 2	Experimental	0.28	22	0.00
	Control	0.30	26	0.00

5.1.4.2 Homogeneity of Variance (Q1a)

To test for homogeneity of variance a Levene's test was conducted. The data meets homogeneity of variance at time 1 and time 3 (Table 5-5).

Table 5-5: A table showing the results of the Levene's test of homogeneity of variance

	Statistic	**Df1**	**Df2**	**Sig**
Time 1 – Number of students with a structured behaviour plan.	2.91	1	46	0.95
Time 3 – Number of students with a structured behaviour plan.	2.35	46	62	0.13

5.1.5 Analysis of Research Question 1a (i)

As the data meets the assumptions of homogeneity of variance, although it does not meet the assumptions of normal distribution, it can be treated as such, therefore the data will be analysed using parametric tests (ANOVAs are fairly robust to violations of assumptions (Pallant, 2013)). The parametric test suited to this data is a two-way mixed ANOVA due to the data having a within group independent variable (time), a between group independent variable (group) and a continuous dependent variable (number of plans) (Pallant, 2013; Field, 2013). For a two-way repeated measures ANOVA to be conducted, data must meet a further parametric assumption of normal covariance distribution. The results of the "Box test" show that this assumption has not been violated (p=0.001).

Table 5-6 and Figure 5-1 show that the number of behaviour plans increased over time for both groups. The experimental group use of behaviour plans is increased from time 1 to time 3 by 1.57, but the control group use of behaviour plans is also increased (by 1.66). Further analysis will determine whether these figures reach statistical significance.

Table 5-6: A table showing the means and standard deviations for the number of students reported to have behaviour plans in place by respondents.

		Mean	**Standard Deviation** (sd)	**Difference** Mean (sd)
Experimental Group (N=21)	Time 1	1.14	1.68	**1.57 (2.36)**
	Time 3	2.71	4.04	
Control Group (N=27)	Time 1 (N=29)	0.67	1.07	**1.66 (0.71)**
	Time 3 (N=35)	2.33	1.88	

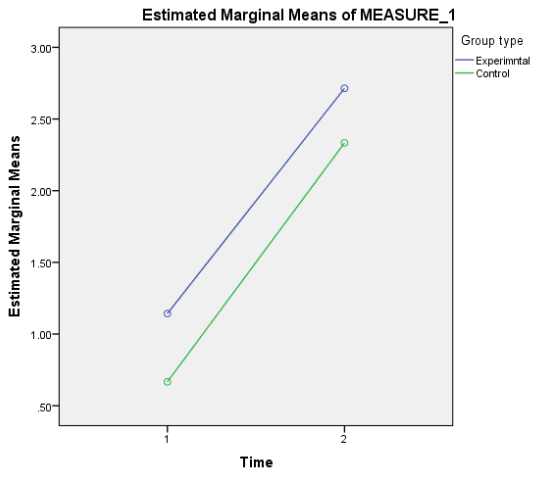


Figure 5-1: A graph showing the number of behaviour plans in place for time 1 and time 3 according to group.

The ANOVA shows that there is a significant main effect of time ($F(1,46)=13.66$, $p=0.001$, partial eta squared=0.23). The results also show no significant interaction of time and group on the number of behaviour plans in place ($F(1,46)=0.01$, $p=0.91$, partial eta squared<0.001). The results show no significant effect of training on the number of behaviour plans in class. Consequently the null hypothesis must be accepted and the experimental hypothesis rejected.

5.1.6 Analysis of Research Question 1a (ii)

Data informing the response to question 1a was requested from participants, but due to its nature cannot be analysed through statistical tests and must therefore be considered through frequencies only. Frequencies are shown in Table 5-7.

Table 5-7: A table showing the number of staff implementing / not implementing the behaviour plan provided in the training and their beliefs about whether this was helpful.

	Yes Number of participants (% of total experimental group)	No Number of participants (% of total experimental group)	Partially Number of participants (% of total experimental group)	Missing Data Number of participants (% of total experimental group)
Have you implemented the behaviour plan?	13 (27.1%)	16 (33.3%)	3 (6.3%)	16 (33%)
Do you feel the behaviour plan was helpful for the student?	13 (27.1%)	1 (2.1%)	2 (4.2%)	32 (66.7%)
Do you hope to implement the behaviour plan?	18 (37.5%)	11 (22.9%)		19 (39.6%)

5.1.7 Summary of Analysis Research Question 1a

The results from this analysis suggest that training did not increase the likelihood of increasing the use of a behaviour plan in the classroom. Although the number of behaviour plans used after training did increase over time, the same occurred (to a greater extent) in the group who did not receive training. Therefore the experimental hypothesis must be rejected.

Further consideration of the results showed that 27% of staff receiving training implemented the behaviour plan after training, and 6% partially implemented the behaviour plan. Of those implementing the behaviour plan, 81% felt it was helpful to the student, 12% were unsure and only 1 participant felt it was not helpful. A further 36% hope to implement the behaviour plan, suggesting that training may motivate implementation.

5.1.8 Research Question 1b and 1c – Analysis and Results

The tests for hypothesis 1b and 1c are the same with different dependent variables (self-efficacy scores, or burnout scores). The pre-tests will be reported together. The independent variables are condition (control, experimental) and time (time 1, time 2, time 3), the dependent variables are self-efficacy and burnout scores.

5.1.8.1 Assumption of Normality (Q1b and Q1c)

Other than 1 actual self-efficacy scale at time 2, all actual self-efficacy scores meet the assumption of normal distribution (Appendix 17). When estimated scores are included, 3 collections of subscale data are not normally distributed. More subscales on the burnout measures do not meet the assumption of normal distribution. However, the majority of the results do. It is unclear if results differ too far from normal distribution for parametric tests given these results below. Pallant (2013) suggests that for samples over 30 parametric tests such as ANOVA's are robust enough to be more appropriate than non-parametric equivalents regardless of normal distribution. Due to the participant numbers further tests will be conducted to determine whether the data is suitable for parametric analysis.

5.1.8.2 Homogeneity of Variance (Q1b and Q1c)

A Levene's test of homogeneity of variance will be presented individually for each hypothesis. This is due to each hypothesis including a different number of participants.

5.1.8.3 Equivalency of Groups (Q1b and Q1c)

In order to assess equivalency of groups at time 1 independent t-tests will be performed on time 1 data.

Self - Efficacy

The Levene's test completed on time 1 data shows homogeneity of variance can be assumed (Table 5-8). The t-test results show there is no significant difference between groups. Therefore groups can be said to be equivalent at time 1 for actual and estimated self-efficacy scores.

Table 5-8: A table showing the Levene's test and equivalency of group t-test results for self-efficacy scores at time 1.

	Levene's test		**Independent Samples t-test**		
	F	p	Df	t	p
Actual Scores					
Self – Efficacy Overall (N=108)	0.09	0.76	53	0.26	0.79
Self-Efficacy Personal (N=108)	1.11	0.29	106	0.86	0.93
Self-Efficacy General (N=59)	0.21	0.65	57	-0.65	0.52
Self-efficacy External (N=55)	0.62	0.43	97	1.05	0.28
Including Estimated Scores (N=108)					
Self – Efficacy Overall	0.16	0.7	104	-1.00	0.32
Self-Efficacy Personal	0.90	0.35	106	-0.58	0.57
Self-Efficacy General	1.24	0.27	105	-1.48	0.14
Self-efficacy External	0.40	0.53	105	0.28	0.78

Burnout

The Levene's test completed on time 1 data shows homogeneity of variance can be assumed (Table 5-9). The t-test results show there is no significant difference between groups. Consequently the groups can be said to be equivalent at time 1 for burnout scores.

Table 5-9: A table showing the Levene's test and equivalency of group t-test results for burnout scores at time 1

	Levene's test		**Independent Samples t-test**		
	F	p	Df	t	p
Actual Scores					
Burnout Overall (N=100)	0.27	0.60	98	1.46	0.15
Burnout Physical (N=103)	1.44	0.23	101	1.41	0.16
Burnout Emotional (N=103)	0.04	0.84	101	0.34	0.73
Burnout Cognitive (N=103)	0.21	0.65	101	1.31	0.19

5.1.9 Analysis of Research Question 1b

1) In a sample of rural primary schools can training based on a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms:

b) Increase overall staff self-efficacy in behaviour management?

Hypotheses

b(i) Participants' self-efficacy scores will be increased after training and this will be maintained 6-8 weeks later.

b(ii) Participants' receiving training will have a greater increase in self-efficacy after training (6-8 weeks) than participants who have not received any training.

5.1.9.1 Analysis of Hypothesis 1b (i)

A two way mixed ANOVA will be conducted to compare self-efficacy scores over time 1, 2, and 3 for the experimental participants. Where sphericity is not assumed multivariate analysis results will be reported as these make more reserved calculations that do not assume sphericity (Pallant, 2013). Where no significant effects are found only overall results will be reported. On finding significant main effects, comparisons between individual times will also be reported.

Personal Self-efficacy

24 participants completed all of the responses for 3 of the questionnaires for personal self-efficacy. The means and standard deviations suggest a slight decrease in personal self-efficacy after training followed by an increase at time 3 (6-8 weeks later) to slightly higher than at time 1 (Table 5-10).

Table 5-10: A table showing the means and standard deviations for personal self-efficacy in the experimental group.

	Time 1 Mean (sd)	**Time 2** Mean (sd)	**Time 3** Mean (sd)
Personal Self-efficacy (Actual) (N=24)	47.06 (4.38)	46.81 (5.58)	48.46 (5.56)

Mauchley's test of sphericity shows sphericity is assumed (df(2)=0.51, p=0.78). The results of the ANOVA show there are no significant effects of time on personal self-efficacy scores (F(2,46)=2.64, p=0.82, partial eta squared=0.10). Consequently the experimental hypothesis must be rejected for personal self-efficacy.

External Self-Efficacy

23 participants completed all questions in the external self-efficacy scales at all 3 time points. The means (Table 5-11) suggest results support the hypothesis with an increase in external self-efficacy scores over time. ANOVA will determine if these results are significant.

Table 5-11: A table showing the means and standard deviations for external self-efficacy in the experimental group.

	Time 1 Mean (sd)	**Time 2** Mean (sd)	**Time 3** Mean (sd)
External Self-efficacy (Actual) (N=23)	38.41 (6.65)	39.37 (7.61)	39.67 (6.74)

For the actual results Mauchley's test of sphericity shows sphericity is not assumed (df(2)=6.33, p=0.04). The results of the multivariate ANOVA show there are no significant effects of time on external self-efficacy scores (Wilks Lambda = 0.94, F(2, 21)= 0.70, p=0.51, partial eta squared=0.06). As results are not significant the experimental hypothesis must be rejected.

General Self-Efficacy

15 participants completed all the questions in the general self-efficacy scores at all 3 time points. The number of participants when including estimated general self-efficacy scores increases to 25. The means and standard deviations are shown in Table 5-12.

Table 5-12: A table showing the means and standard deviations of general self-efficacy scores in the experimental group.

	Time 1 Mean (sd)	**Time 2** Mean (sd)	**Time 3** Mean (sd)
General Self-efficacy (Actual) (N=15)	60.13(4.45)	59.87 (5.63)	63.13 (7.11)
General Self-efficacy (Estimated) (N=25)	58.57(6.16)	64.50(5.93)	61.62(8.00)

For the actual results Mauchley's test of sphericity shows sphericity is assumed (df(2)=3.12, p=0.80). The results of the ANOVA show no significant effects of time on general self-efficacy scores (F(2,28)=2.52,p=0.09, partial eta Squared=0.15). These results suggest a trend but are not statistically significant so do not support the experimental hypothesis.

Sphericity can be assumed for the data including the estimated values (df(2)=1.6,

p=0.44). The ANOVA shows a significant effect of time on general self-efficacy ($F(2,48)=13.18, p<0.01$, partial eta squared=0.36). Bonferroni pairwise comparisons show the scores in time 1 to be significantly lower than scores at time 2 ($p<0.001$), and time 3 ($p<0.03$). There is no significant difference between time 2 and time 3 ($p>0.05$). These results suggest that general self-efficacy increased significantly from time 1 to time 2, and was then maintained at this level. These results support the experimental hypothesis for general self-efficacy and reject the null hypothesis.

Overall Self-efficacy.

15 participants completed all the questions in the overall self-efficacy scores at all 3 time points. Including estimated self-efficacy scores increases the number of participants to 25. The means and standard deviations are shown in Table 5-13.

Table 5-13: A table showing the means and standard deviations overall self-efficacy in the experimental group.

	Time 1 Mean (sd)	**Time 2** Mean (sd)	**Time 3** Mean (sd)
Overall Self-efficacy (Actual) (N=15)	146.13(12.30)	148.20 (7.69)	152.70 (13.20)
Overall Self-efficacy (Estimated) (N=25)	143.89 (12.83)	149.96 (13.60)	149.84 (14.59)

For the actual results Mauchley's test of sphericity shows sphericity is assumed ($df(2)=4.25$, $p=0.12$). The results of the ANOVA show there is a significant effect of time on overall self-efficacy scores ($F(2,28)=3.29, p=0.05$, partial eta Squared = 0.19). Bonferroni pairwise comparisons show no significant effects between individual time periods; this suggests scores underwent a significant gradual increase over time. Actual data supports the experimental hypothesis and rejects the null hypothesis.

For results including the estimated values sphericity is assumed ($df(2)=2.89$, $p=0.24$). The ANOVA shows a significant effect of time on overall self-efficacy ($F(2,48)=7.30, p<0.005$, partial eta squared= 0.23). Bonferroni comparisons show time 2 is significantly higher than time 1 ($p = 0.001$), as is time 3 ($p<0.03$). There is no

significant difference between time 2 and 3, suggesting the change in overall self-efficacy scores after time 1 is maintained until follow up (when including estimated scores). The results including estimated scores also support the experimental hypothesis.

5.1.9.2 Analysis of Hypothesis 1b (ii)

The data meets parametric assumptions for homogeneity of variance and although it does not consistently meet assumptions of normal distribution for all dependent variables the sample size is sufficient to allow for parametric testing of the data (5.1.8.1). Two way repeated measures ANOVAs will be conducted for hypothesis 1b(ii). The independent variables are group (experimental and control) and time (time 1 and time 3), the dependent variables are the self-efficacy scores.

Two way repeated measures ANOVAs require equality of covariance, tested by Box's test of equivalence. To meet the criteria of equality of covariance the box test must be >0.001. All analyses showed this result of the Box's test to be non-significant. Therefore the two way repeated measures ANOVA was appropriate for all dependent variables.

Personal Self-Efficacy

Means and standard deviations for actual results are included in Table 5-14. Due to the small amount of missing data accounted for by the estimated values only the actual scores will be used for analysis.

Table 5-14: A table showing the means and standard deviations for personal self-efficacy scores for both groups at time 1 and time 3.

		Time 1 Mean (sd)	**Time 3** Mean (sd)
Personal Self-efficacy (Actual)	Experimental (N = 33)	47.83 (5.35)	50.24 (4.12)
	Control (N= 35)	49.09 (5.87)	48.86 (4.87)

The results showed a significant interaction for personal self-efficacy scores between time and group (Wilks Lambda=0.93, $F(1,66)=5.38$, $p<0.03$, partial eta squared=0.08). There was no main effect of time on scores (Wilks Lambda=1, $F(1,66)=0.01$, $p>0.05$, partial eta squared <0.005). This shows a significant difference in scores according to group, this supports the experimental hypothesis.

External Self-Efficacy

Means and standard deviations for external self-efficacy scores are included in Table 5-15 and suggest results support the hypotheses.

Table 5-15: A table showing the means and standard deviations for external self-efficacy scores for both groups at time 1 and time 3.

		Time 1 Mean (sd)	**Time 3** Mean (sd)
External Self-efficacy (Actual)	Experimental (N = 32)	39.77(6.98)	39.89(6.98)
	Control (N= 35)	38.77(7.53)	38.30(8.42)

The results for external self-efficacy show no significant interaction between group and time (Wilks Lambda=1.00, $F(1,65) = 0.11$, $p=0.75$, partial eta squared<0.001). The results also show no significant effect of time (Wilks Lambda = 1.00, $F(1,65) = 0.04$, $p=0.85$, partial eta squared=0.001). This analysis supports a rejection of the hypothesis for external self-efficacy.

General Self-Efficacy

General self-efficacy data was markedly impacted by missing values. Consequently both actual scores and scores including estimated results will be analysed. Descriptive statistics (Table 5-16) suggest support for the experimental hypotheses with both group general self-efficacy scores increasing over time, but with a greater increase in the experimental group.

Table 5-16: A table showing the means and standard deviations for general self-efficacy scores for both groups at time 1 and time 3.

		Time 1 Mean (sd)	**Time 3** Mean (sd)
General Self-efficacy (Actual)	Experimental (N = 22)	59.14 (5.24)	63.23 (6.38)
	Control (N= 18)	59.89 (4.19)	60.83 (3.50)
General Self-efficacy (Estimated)	Experimental (N = 35)	58.26 (5.92)	62.31(6.95)
	Control (N= 37)	59.16(5.66)	61.84(5.25)

The ANOVA for actual data scores show no significant interaction (Wilks Lambda=0.93, F(1,38) =3.01, p=0.09, partial eta squared<0.001). A significant main effect of time is found (Wilks Lambda=0.83, F(1, 38), p<0.01, partial eta squared=0.17).

The ANOVA for scores including estimated scores also shows no significant interaction (Wilks Lambda=0.96, F(1,70)=2.73, p=0.10, partial eta squared<0.04). A main effect of time was found (Wilks Lambda=0.78, F(1,70)=19.27, p<0.001, partial eta squared=0.22). Both estimated and actual data scores reject the hypothesis showing general self-efficacy scores increase over time with no significant impact of training on scores.

Overall Self-Efficacy

Overall self-efficacy scores were considerably impacted by missing data values on individual items. Consequently data will be analysed for actual and estimated scores. Descriptive statistics (Table 5-17) show an increase in scores for the experimental group over time, with a decrease in scores for the control group. An interaction also takes place for actual scores and scores including estimated scores.

Table 5-17: A table showing the means and standard deviations for overall self-efficacy scores for both groups at time 1 and time 3.

		Time 1 Mean (sd)	**Time 3** Mean (sd)
Overall Self-efficacy (Actual)	Experimental (N = 22)	147.91(13.38)	152.70(14.04)
	Control (N= 17)	148.91(12.21)	145.71(12.30)
Overall Self-efficacy (Estimated)	Experimental (N = 35)	145.26 (13.28)	150.84 (15.14)
	Control (N= 37)	149.31 (10.80)	148.61 (14.27)

The ANOVA for actual data scores shows a significant interaction between experimental and control group overall self-efficacy scores (Wilks Lambda=0.10, F(1,37)=3.95, p=0.05, partial eta squared<0.10). There is no significant main effect of time (Wilks Lambda=0.10, F(1, 37), p=0.70, partial eta squared=0.04). The ANOVA results, with included estimated values, show a significant interaction between control and

experimental overall self-efficacy scores (Wilks Lambda=0.93, F(1,70) =5.04, p<0.03, partial eta squared=0.07). No significant main effect of time was found (Wilks Lambda=0.96, F(1,70)=3.06, p=0.09, partial eta squared<0.04). These results accept the experimental hypothesis.

5.1.10 Summary of Results - Research Question 1b

The results show that the effect of training appears to have a different impact on the different subscales of self-efficacy. The results also suggest that staff self-efficacy scores are not constant over time.

Results for hypothesis 1b(i) show no impact of training on personal and external self-efficacy scores (with small to medium effect sizes). While the actual data taken for general self-efficacy showed no significant effect of training, the means suggested a trend toward supporting the hypothesis. When values analysed included the estimated scores, a highly significant effect was found ($F(2,48)=13.18$, $p<0.001$) with a large effect size. The results showed a significant increase after time 1 to time 2 ($p<0.001$); this increase was maintained at time 3 ($p<0.03$). The overall self-efficacy actual scores showed a significant increase in results over time ($F(2,28)=3.29, p=0.05$) with a large effect size. When estimated data was included this effect was increased ($F(2,28)=7.3, p<0.005$) as was the effect size. The estimated data showed a significant increase from time 1 to time 2 ($p<0.005$), and time 1 to time 3 ($p<0.03$). The results support acceptance of the null hypothesis for personal and external self-efficacy scores. The experimental hypothesis can be accepted for general (when estimated scores are included) and overall (based on actual and estimated scores) self-efficacy scores.

Analysis was also conducted comparing the experimental and control group scores before training and at follow up (1b(ii)). The experimental hypothesis must be rejected for external self-efficacy and general self-efficacy scores (although a non-significant trend supporting the hypothesis is suggested for external self-efficacy scores). The experimental hypothesis can be accepted for personal self-efficacy ($F(1,66=5.38, p<0.03$) with a medium to large effect size. The experimental hypothesis can also be accepted for the impact of training on overall self-efficacy scores ($F(1,37)=3.95, p=0.05$) (actual), ($F(1,70)=5.04, p<0.03$) (estimated)) with a medium to large effect size.

5.1.11 Analysis of Research Question 1c

1) In a sample of rural primary schools can training based on a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms:

c) Decrease teacher burnout?

Hypotheses

c(i) Participants' burnout scores will be decreased after training and this will be maintained 6-8 weeks later.

c(ii) Participants' receiving training will have a greater decrease in burnout scores after training (6-8 weeks later) than participants who have not received any training.

5.1.11.1 Analysis of Hypothesis 1c (i)

A two way mixed ANOVA will be used to compare burnout scores over time 1, 2, and 3 for the experimental participants. Where sphericity is not assumed multivariate analysis results will be reported (these do not assume sphericity). Where no significant effects are found, only overall results will be reported. Where significant effects are found Bonferroni comparisons between individual times will also be reported.

Physical Fatigue

The means and standard deviations for physical burnout scores are shown in Table 5-18 and suggest support for the hypothesis with a decrease in burnout over time.

Table 5-18: A table showing the means and standard deviations physical burnout in the experimental group.

	Time 1 Mean (sd)	**Time 2** Mean (sd)	**Time 3** Mean (sd)
Physical Burnout (Actual) (N=26)	3.38(0.96)	3.36 (1.19)	2.91 (1.16)

Mauchley's test of sphericity shows sphericity is not assumed (df(2)=7.90, p=0.02). The results of the multivariate ANOVA show there is no significant effect of time on physical burnout (Wilks Lambda=0.81,F(2,24)=2.75,p=0.08, partial eta squared=0.19). This suggests the null hypothesis must be accepted for physical burnout.

Emotional Exhaustion

24 participants completed all the questions in the emotional burnout scales at all 3 time points. The means and standard deviations are shown in Table 5-19 and suggest an effect opposite to that predicted by the hypothesis.

Table 5-19: A table showing the means and standard deviations emotional burnout in the experimental group.

	Time 1 Mean (sd)	**Time 2** Mean (sd)	**Time 3** Mean (sd)
Emotional Burnout (Actual) (N=24)	1.85(1.32)	2.78 (0.88)	2.10 (1.09)

Mauchley's test of sphericity shows sphericity is assumed (df(2)=1.76, p=0.42). The results of the ANOVA show a significant effect of time on emotional burnout (F(2,46)=4.83,p=0.01, partial eta squared=0.17). Bonferroni pairwise comparisons show a significant increase in emotional burnout scores between time 1 and time 2 ($p<0.04$), and a significant decrease in emotional burnout scores between time 2 and time 3 (p=0.05). There is no significant difference between time 1 and time 3 scores. The experimental hypothesis must be rejected.

Cognitive Weariness

25 participants completed all the questions in the cognitive burnout scores at all 3 time points. Means and standard deviations (Table 5-20) show the immediate burnout scores decreasing after training but then increase again for follow up.

Table 5-20: A table showing the means and standard deviations cognitive burnout in the experimental group.

	Time 1 Mean (sd)	**Time 2** Mean (sd)	**Time 3** Mean (sd)
Cognitive Burnout (Actual) (N= 25)	2.77(0.81)	1.63 (0.74)	2.60 (0.64)

Mauchley's test of sphericity shows sphericity is assumed (df(2)=5.42, p=0.07). The results of the ANOVA show there is a significant effect of time on burnout scores (F(2,48)=28.27,$p<0.001$, partial eta squared=0.54). Bonferroni pairwise comparisons show a significant decrease in cognitive burnout scores between time 1 and time 2 ($p<0.001$) and a significant increase in cognitive burnout between time 2 and time 3 ($p<0.001$). No significant difference can be seen between time 1 and time 3.

Overall Burnout

23 participants completed all the questions in the overall burnout scores at all 3 time points. The means and standard deviations are shown in Table 5-21, these suggest a decrease in burnout scores after training.

Table 5-21: A table showing the means and standard deviations of overall burnout in the experimental group.

	Time 1 Mean (sd)	**Time 2** Mean (sd)	**Time 3** Mean (sd)
Overall Burnout (Actual) (N=23)	2.67(0.68)	2.63 (0.70)	2.64 (0.70)

Mauchley's test of sphericity shows sphericity is assumed (df(2)= 2.42, p=0.30). The ANOVA shows no significant effects of time on overall burnout scores (F(2,44)=0.08,p=0.92, partial eta squared=0.01). Consequently the null hypothesis must be accepted for overall burnout scores.

5.1.11.2 Analysis of Hypothesis 1c (ii)

Analysis of data will undertake the same procedure as for hypothesis 1b (ii) (see 5.1.9.2 for details). Results for burnout were not as impacted by missing item responses as self-efficacy. Consequently only data from actual responses will be analysed.

Physical Fatigue

Physical burnout score descriptives suggest physical burnout decreased over time, with a greater decrease for the experimental group (Table 5-22). This suggests support for the experimental hypothesis, ANOVA will determine any significance.

Table 5-22: A table showing the means and standard deviations for physical burnout scores for both groups at time 1 and time 3.

		Time 1 Mean (sd)	**Time 3** Mean (sd)
Physical Burnout	Experimental (N= 34)	3.63 (1.16)	3.01 (0.94)
	Control (N= 34)	2.90 (1.17)	3.00 (1.21)

The ANOVA for actual data scores shows a significant interaction between group and time on physical burnout scores (Wilks Lambda=0.89, $F(1,66)= 8.43$, $p<0.01$, partial eta squared=0.11). There is a significant main effect of time (Wilks Lambda=0.94, $F(1,66)=4.14$, $p<0.05$, partial eta squared=0.06). These results support the hypothesis, rejecting the null hypothesis.

Emotional Exhaustion

The means for emotional burnout scores show emotional burnout for both groups increases over time with a greater increase in the control group (Table 5-23). This may suggest support for the experimental hypothesis if the difference is significant.

Table 5-23: A table showing the means and standard deviations for emotional burnout scores for both groups at time 1 and time 3.

		Time 1 Mean (sd)	**Time 3** Mean (sd)
Emotional Burnout	Experimental (N= 34)	1.90 (1.13)	2.15 (2.00)
	Control (N= 31)	1.63 (0.77)	2.10 (0.86)

The ANOVA for actual data scores shows no significant interaction (Wilks Lambda=0.99, F(1,63)=0.51, p=0.58, partial eta squared=0.01). A significant main effect of time on emotional burnout scores was found (Wilks Lambda=0.92, F(1,63),p<0.03, partial eta squared=0.08), showing a significant increase in emotional burnout over time. These results reject the hypothesis and accept the null hypothesis.

Cognitive Weariness

The means for cognitive burnout scores decrease over time in the experimental group and increase over time in the control group, this suggests support for the hypothesis (Table 5-24).

Table 5-24: A table showing the means and standard deviations for cognitive burnout scores for both groups at time 1 and time 3.

		Time 1 Mean (sd)	**Time 3** Mean (sd)
Cognitive Burnout	Experimental (N= 35)	2.78 (1.13)	2.54 (0.78)
	Control (N= 34)	2.36 (0.79)	2.59 (0.92)

The ANOVA for actual data scores shows a significant interaction between time and group (Wilks Lambda = 0.94, F(1,67) = 4.19, p<0.05, partial eta squared = 0.06). There was no main effect of time on cognitive burnout scores (Wilks Lambda = 1.00,F(1,67), p=0.95, partial eta squared<0.01). These results support the hypothesis, rejecting the null hypothesis.

Overall Burnout

The means for overall burnout scores suggest results support the hypothesis. Scores decrease over time in the experimental group and increase over time in the control

group (Table 5-25).

Table 5-25: A table showing the means and standard deviations for overall burnout scores for both groups at time 1 and time 3.

		Time 1 Mean (sd)	**Time 3** Mean (sd)
Overall Burnout	Experimental (N= 33)	2.77 (1.02)	2.58 (0.81)
	Control (N= 34)	2.32 (0.82)	2.57 (0.90)

The ANOVA for overall burnout scores shows a significant interaction between time and group (Wilks Lambda=0.94, $F(1,65) = 4.47$, $p<0.04$, partial eta squared=0.06). There was no main effect of time on overall burnout scores (Wilks Lambda = 1.00, $F(1, 65)$, $p=0.85$, partial eta squared<0.01). These results support the hypothesis, rejecting the null hypothesis.

5.1.12 Summary of Results – Research Question 1c

The results suggest that training appeared to impact individual subscales of burnout differently. For hypothesis 1c(i) the experimental hypothesis must be rejected for overall and physical burnout scores, both of these suggest a trend in the means supporting the hypothesis but this does not reach significance. The results in emotional burnout suggest that the training immediately increases emotional burnout in staff ($p<0.04$), opposite to the effect the hypothesis predicts, although this increased burnout is not maintained over time the hypothesis must be rejected for emotional burnout. Although the cognitive burnout scores do decrease post-training ($F(2,48)=28.27, p<0.01$), as this effect is not maintained over time the experimental hypothesis must also be rejected. Hypothesis 1c(i) must be rejected for all measures of burnout. Training does not appear to significantly reduce teacher burnout measures over time.

Hypothesis 1c(ii) must be rejected for emotional burnout as there is no significant effect of group on scores over time. Emotional burnout significantly increases in both groups over time ($F(1,66)=4.14, p<0.03$), although this increase is less in the experimental group this difference is not significant. The experimental hypothesis can be accepted for physical burnout with a significantly greater decrease over time in the experimental group ($F(1,66)=8.43, p<0.01$) with a moderate to large effect size. The experimental hypothesis can also be accepted for cognitive ($F(1,67)=4.19, p<0.05$) and overall burnout ($F(1,65)=4.47, p<0.04$) for both of these subscales the experimental group burnout is decreased and the control group is increased with moderate effect sizes. Training appears to have had a protective effect for teacher burnout, by reducing burnout, or minimising the increase of burnout over time for physical, cognitive, and overall staff burnout.

5.2 Research Question 2

This section will detail how the results will be analysed for Question 2. Following this each case will be presented individually.

Research Question 2: Does use of a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms decrease challenging behaviour in students?

Hypothesis:

Participant challenging behaviour will decrease over time after implementation of the behaviour plan.

5.2.1 Data Analysis for SCED

The most common method of analysis for SCED data is visual analysis. Although a number of researchers question the accuracy of this approach (Matyas & Greenwood, 1990) this is still the most dominant method of SCED analysis within the literature (Kratochwill et al, 2010). A number of researchers argue for a variety of statistical tests to add validity and reliability to SCED analysis (Palmer, 2010). These include multi-level models, non-parametric tests, and regression models (Palmer, 2010). A number of recommendations regarding visual analysis as well as discussions of various statistical analysis methods are discussed in the literature and will be summarised below.

5.2.1.1 Visual Analysis

Visual analysis is the traditional and most common method of analysis in SCED literature (Kratochwill et al, 2010; Matyas & Greenwood, 1990). Kazdin (2003) states that the aims of visual analysis are to identify any consistent and reliable effects which are likely to be due to the independent variable. Visual analysis has been critiqued by researchers in the field as not as objective as statistical tests that consider explicit statistical criteria to determine impact of an independent variable (Kazdin, 2003).

Traditionally, visual analysis is how SCED data has been analysed, due to the ability of graphical representation of data to show relationships over time, and visual analysis not being subject to "arbitrary" mathematical or statistical assumptions (Cooper et al, 2007). Cooper et al (2007) support the visual analysis of data on the basis that it is a "conservative method" (p128) compared to statistical tests, as the data must also look significant when plotted and not rely only on mathematical probabilities. Visual analysis is said to encourage researchers to make independent judgements and interpret the meaning of behavioural change, providing accessible feedback to those involved in data collection (Cooper et al, 2007).

Matyas & Greenwood (1990) remind researchers of the risk of human error biasing results in their research paper. Matyas & Greenwood (1990) asked undergraduate students to visually analyse a number of graphs and showed that visual analysis increases the likelihood of type 2 errors. Kratochwill et al (2010) have since recommended a number of criteria for visual analysis to support the demonstration of relationships between the independent and dependent variables (such as inclusion of trend and variability lines). Baer (1977) argues that although visual analysis has an increased risk of type 2 error there is also a concurrent decreased risk of type 1 errors compared to statistical techniques. Baer (1977) argues that smaller, less consistent differences will be validated through statistical techniques than would be through visual analysis. Visual analysis is argued to require larger more consistent changes to be observed than would be required to be found significant in statistical testing methods, leading to "more powerful, general, dependable" and "actionable" conclusions (Baer, 1977 p.171).

Kratochwill et al's (2010) technical standards document for SCEDs includes a number of recommendations for visual analysis (Table 5-26). They report that visual analysis should always give consideration to the predictability of the baseline, data within each phase independently, and between phases, then integrate these to inform conclusions. This paper reports that for an effect to be described as "strong" the analysis needs to show 3 demonstrations of effect, and 0 demonstrations of "non-effect" (see Table 5-27 for definitions). 3 demonstrations of effect and 1 demonstration of non-effect would be

sufficient to be defined as “moderate effect” and less than 3 demonstrations of effect would be defined as “non-effect”.

Table 5-26: A table to show the features for visual analysis (adapted from Kratochwill et al, 2010).

Feature	**Definition**
Level	The mean score for data within a phase.
Trend	The slope of the best fitting line within each phase.
Variability	The range of standard deviation around the trend line.
Immediacy of effect	The change between the last three data points in one phase and the first three data points in the next phase.
Overlap	Proportion of data from one phase that overlaps from a previous phase.
Consistency of data patterns across similarity of phases	Consistency of data across similar phases (relevant for multiple baseline or withdrawal type designs).

Table 5-27: A table of "non-effects" in SCED visual analysis (adapted from Kratochwill et al, 2010).

Potential area for non - effect	**Definition**
Baseline	A failure to show a sufficient demonstration of pattern of responding.
Within Phase	Failure to establish a pattern of responses within any phase.
Across Phases	Inconsistent patterns across similar phases (for withdrawal or multiple baseline SCED's only).
Across Phases (Overlap)	A long latency between introduction of the independent variable and a change to the outcome variable. (Overlap between observed and projected patterns of the dependent variable).

Due to the consistent use of visual analysis in SCED research, evidence that this analysis method can lead to valid and useful conclusions (Kratochwill et al, 2010; Cooper at el, 2007; Baer et al, 1977), and failure to identify mathematical analyses that can adequately replace visual analysis (see 5.2.1.2), the current research will use visual analysis. In order to decrease risk of human error the standards from Kratochwill et al (2010) will guide the visual analysis process, including the recommendations for ensuring inter-observer reliability before asserting conclusions.

5.2.1.2 Statistical Analysis

Due to reported inaccuracies of visual analysis (Matyas & Greenwood, 1990; Baer, 1977) there has been an increasing emphasis on working towards development and identification of quantitative SCED analysis methods (Parker, Vannest & Davis, 2011). A number of quantitative measures focused on determining effect sizes (Kratochwill et al, 2010) have been considered, including methods based on numerical calculations around overlapping and non-overlapping data between phases, and regression estimators. This discussion will focus around literature that discusses the types of analyses within the SCED literature. Due to the developing and fairly recent area of research there are a wide range of methods undergoing testing. For the purpose of this review the most commonly tested and reviewed methods will be discussed.

Barlow et al's (2009) and Barlow & Hersen's (1984) books consider a number of viewpoints on the use of statistical analysis methods for SCED research, highlighting the ongoing debate surrounding the usefulness of statistical analyses with authors questioning the value of added statistical testing procedures. Kazdin (1984) argues that statistical methods should be seen as providing potential supplementary information and not as a replacement to visual analysis, especially in cases where patterns in data are not apparent in visual analysis. However, arguments presented by researchers such as Baer (1977) (see 5.2.1.1.) would suggest statistical methods increase the risk of type 1 error which is minimised through visual analysis procedures.

There are risks of using statistical analyses due to the risk of tests being insensitive to outliers (Parker et al, 2011) and their lack of applicability where autocorrelation and serial dependence apply to the data properties (Barlow et al, 2009). Research has shown that even very low levels of autocorrelation can have a large impact on statistical test analysis (Manolov & Solanas, 2008). Due to on-going research into the area of SCED statistical analysis (Parker et al, 2011) a number of statistical procedures have been explored to determine their applicability and appropriateness for the current research.

Use of T and F Tests

As these tests are commonly used and were developed for comparing differences between groups, while ignoring trends within the data (Barlow et al, 2009), they are often not seen as useful for SCEDs. These tests are not applicable to SCEDs as they require independent data and normal distribution (Barlow et al, 2009). Some researchers use randomisation tests when data is not normally distributed. However this SCED did not use random allocation to treatment or measurement and therefore these tests are not deemed appropriate.

Effect Size Calculations

Due to the calculations developed for effect size being developed for group designs (using average levels from baseline and treatments) it has been argued that these methods are inappropriate for a SCED (Manolov & Solanas, 2008). Ross (2012) identified 3 calculations of effect size; regression, percentage of non-overlap (see below), and standardized mean difference. Manolov & Solanas (2008) report that these tests are very susceptible to effects of autocorrelation, potentially increasing risk of type 1 errors. Regression analysis is believed to be helpful as they can model trends found within the data (Kratochwill et al, 2010), and consider slope and mean levels (Manolov & Solanas, 2008). However, Manolov & Solanas (2008) found that simpler methods were much more reliable and urge further research into regression analyses for SCED before relying on these. Consequently these methods of effect size calculations are not considered appropriate for the current research.

Percentage of non-overlapping data.

This is the most commonly used statistical analysis method in SCED research (Manolov & Solanas, 2008) aiming to calculate an effect size. This is used within visual analysis and calculates the percentage of data points that do not overlap between phases. Parker et al (2011) report that non-overlap indices are the most robust of the statistical analyses procedures available. Kratochwill et al (2010) reports a number of limitations to this method, such as it being significantly affected by outliers, and not being applicable if the baseline is unstable. This test was found to be one of the best performing tests for decreased impact of autocorrelation and serial dependence on results, however baseline

stability was found to impact the applicability of this method (Manolov & Solanas, 2008).

Interrupted Time Series Analysis.

This technique compares data over time and was designed for single or group cases. The technique aims to control for autocorrelation and parametric data is not required (Barlow et al, 2009). The technique accounts for serial dependency and considers the change in slope and level between phases. The technique requires a minimum of 50 data points, "preferably 100" (Kazdin, 1984, p301), and is therefore not appropriate for this research.

5.2.2 Proposed Data Analysis Procedure

Due to the difficulties with statistical analysis presented above, this research will use visual analysis for research question 2, guided by the standards in Kratochwill et al (2010). Graphs will be presented for each behaviour and then visually analysed for variability, trend, level, overlapping data points, and immediacy of effect.

A brief reminder of the pupil profile and recorded behaviours will be provided, followed by the presentation of the numerical and graph data for each case. A summary highlighting the key observations and behavioural patterns over time will follow; this description will include interpretation of the data in relation to the research question. Discussion of these interpretations can be found in the final chapter.

5.2.3 Inter-rater Agreement

Kratochwill et al (2010) reports visual analysis should include calculation of inter-rater agreement. This was conducted by a colleague of the researcher, who is familiar with SCED graphs. Agreement was calculated using Cohen's Kappa, which is most appropriate for nominal data (Gisev et al, 2013) to calculate agreement for visual analysis. A Cohen's Kappa of 0.4-0.6 indicates a moderate agreement, 0.6-0.8 indicates a substantial agreement, and 0.8 or above an almost perfect agreement (Gisev et al, 2013).

Inter-rater reliability calculations were completed using a scale of 1 to 5. The researcher and rater assigned a value to the statement:

How certain are you that the behaviour decreases markedly after and during the implementation of the behaviour plan?

The values were given labels which were:

1 – not at all, 2 – unsure, 3 – it is possible, 4 – reasonably certain, 5 – very certain.

The Cohens Kappa value was 0.827, indicating a high level of agreement.

5.2.4 Case 1 (S1a)

Name: S1a

Age: 10 years and 1 month years at the beginning of the intervention

Year Group: Year 5

Behaviours Recorded:

Behaviour 1: Inappropriate reaction to not being chosen when his hand is up.

Definition: Responds with verbal displeasure when not chosen if his hand is up (see Appendix 7 for full definition).

Behaviour 2: Interrupting Teacher.

Definition: S1a interrupts the teacher when she is talking to another student or the whole class (see Appendix 7 for full definition).

Behaviour 3: Complaints about lack of help given from staff.

Definition: S1a complains about a lack of support when he is able to complete the work independently (see Appendix 7 for full definition).

Behaviour 4: Refusal to complete work with support.

Definition: S1a refuses to complete work when supported by staff (see Appendix 7 for full definition).

5.2.4.1 S1a – Behaviour 1 (Inappropriate reaction to not being chosen)

Data is presented graphically with key numerical information presented in Table 5-28.

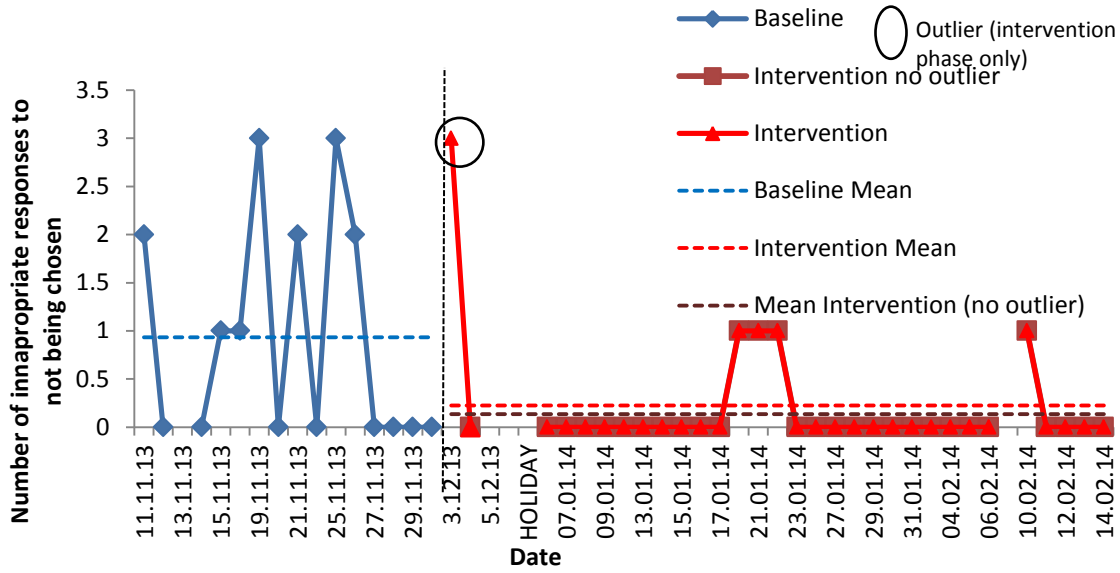


Figure 5-2 A graph showing the number of inappropriate responses per day and level lines for participant S1a across Phase A and B. (Including level line for data not including the outlier in phase B).

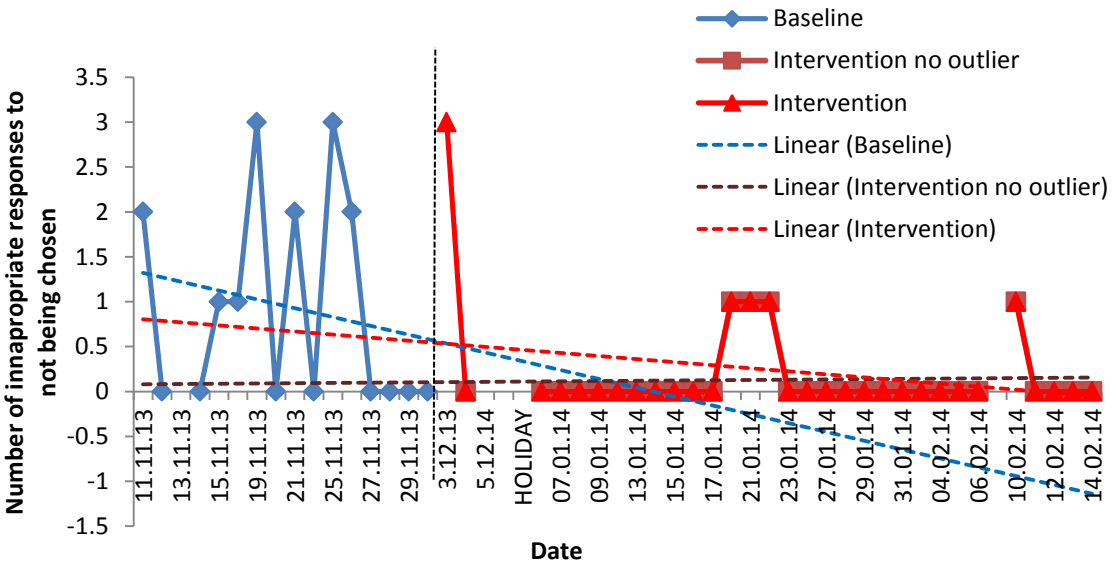


Figure 5-3: A graph showing the number of inappropriate responses per day and trend lines for participant S1a across Phase A and B. (Includes additional trend line not including outlier in phase B).

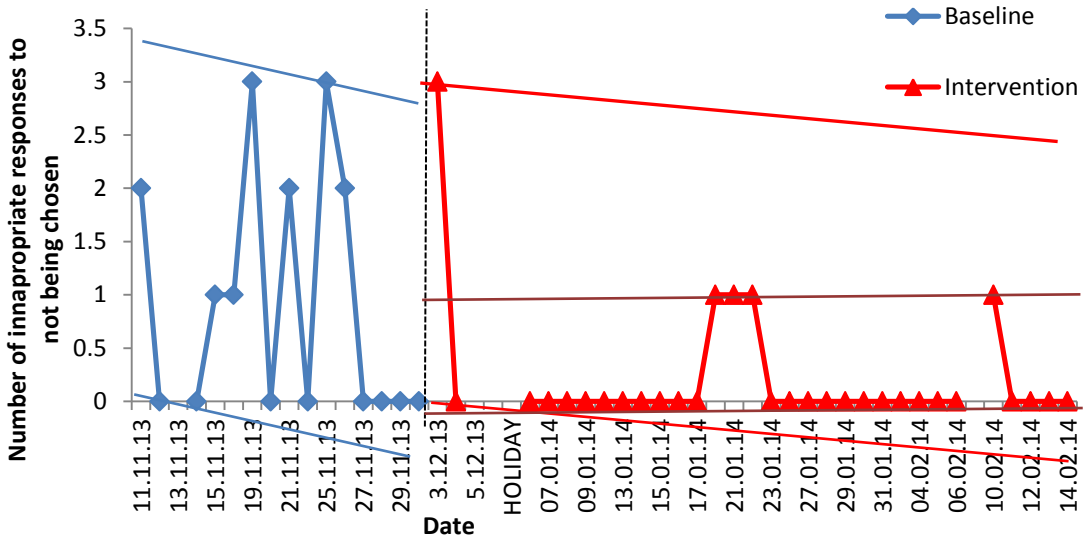


Figure 5-4: A graph showing the number of inappropriate responses per day with variability lines for participant S1a across Phase A and B. An additional variability line has been added which does not include the outlier in the intervention phase.

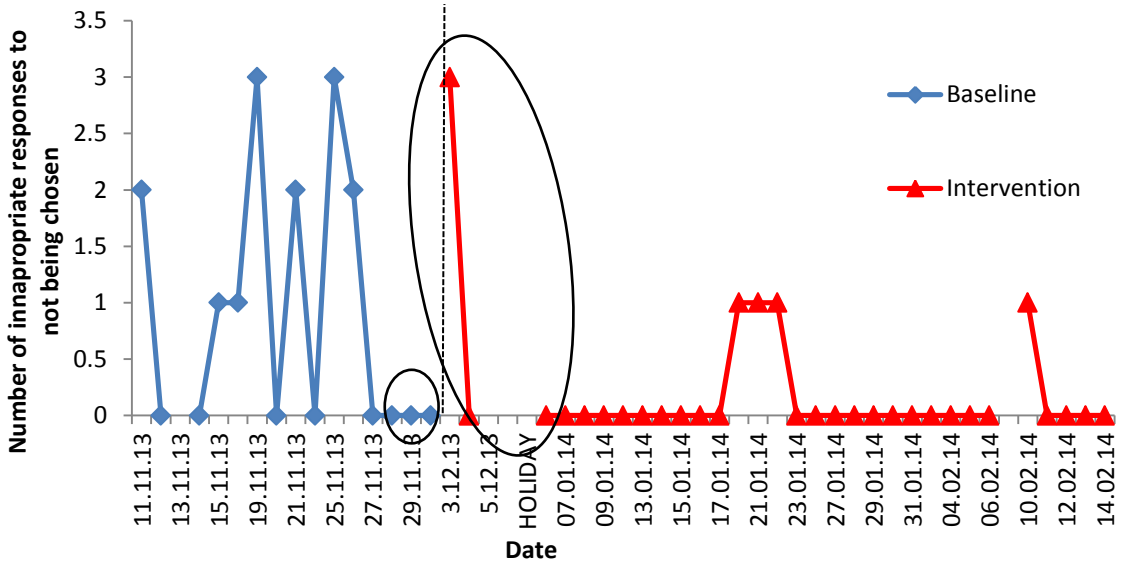


Figure 5-5: A graph showing the number of inappropriate responses per day with immediacy of effect for participant S1a across Phase A and B.

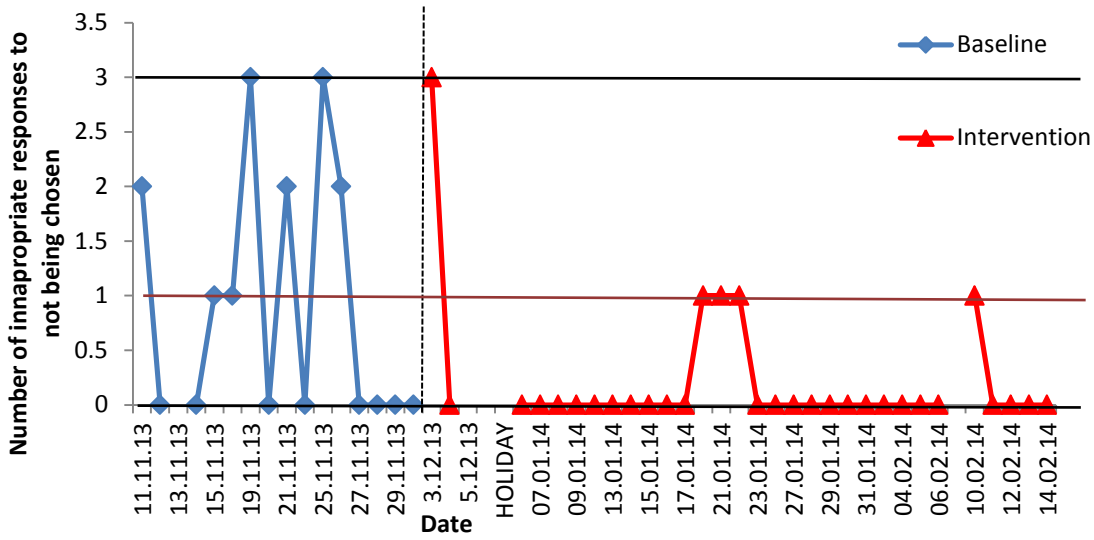


Figure 5-6: A graph showing the number of inappropriate responses per day with percentage of overlap lines for participant S1a across Phase A and B. An additional line has been added which does not include the outlier in the intervention phase.

Table 5-28: A table showing the mean and standard deviation for S1a (number of inappropriate responses per day)

	Baseline	**Intervention**	**Intervention (no outlier)**	**Difference**
Mean	0.93	0.23	(0.13)	0.7 (0.8)
Standard Deviation	1.12	0.61	(0.34)	0.51 (0.78)

Table 5-29: A table outlining brief descriptions of the outcome of the visual analysis of number of inappropriate responses for not being chosen for S1a.

Feature	**Visual Analysis**
Level	Figure 5-2 shows a notable decrease in the mean level (reducing by 0.7 (Table 5-28)). This decrease is made greater (0.8) when the outlier is not included in the analysis.
Trend	Figure 5-3 shows that the trend for the baseline phase was reducing at a quicker rate than the trend for the intervention phase. The intervention trend is also shown without the outlier, which suggests a stable, lower trend line with less of a gradient than the baseline phase.
Variability	The variability lines (Figure 5-4) show that the variance is approximately equal in both phases. A third variance line has been added that does not include the outlier (Figure 5-4) which shows a much smaller variance in the intervention than the baseline phase.
Immediacy of Effect	The intervention immediacy of effect shows an immediate sudden increase in number of inappropriate responses (the outlier) (Figure 5-5). This is then followed by an immediate return to the same level as the last 3 points of the baseline (0).
Overlap	Figure 5-6 shows that both phases overlap one another for all data points. However an additional line has been added which does not include the outlier in the intervention phase. Excluding the outlier in the intervention phase, both the intervention and baseline phase go down to 0, however the baseline is higher than the intervention phase for 5 of the 15 (33%) baseline data points.

5.2.4.2 S1a – Behaviour 2 (Interrupting Teacher)

Data is presented graphically with key numerical information presented in Table 5-28.

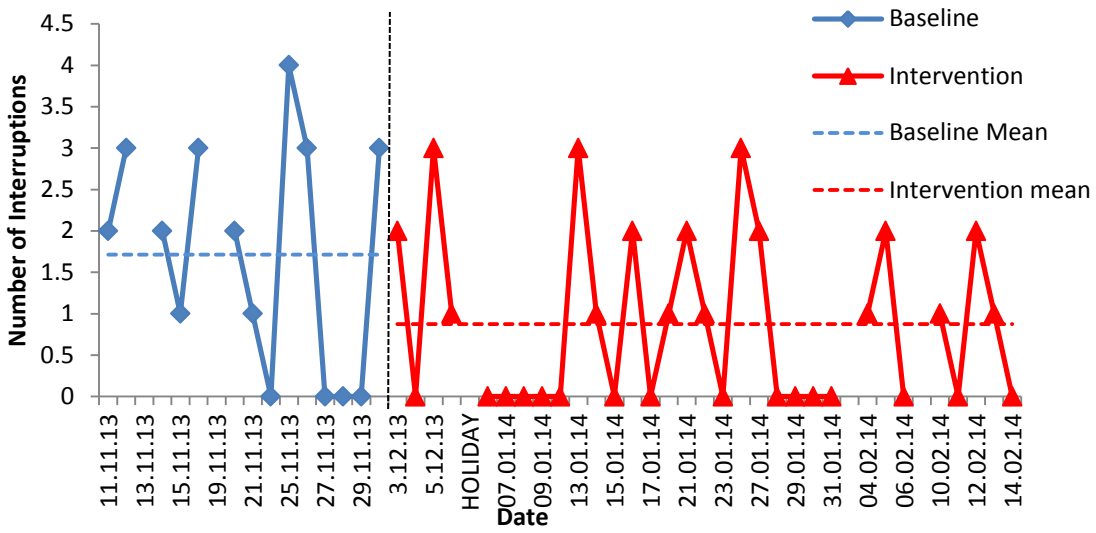


Figure 5-7: A graph showing the number of interruptions of the teacher made by S1a each day, including mean lines.

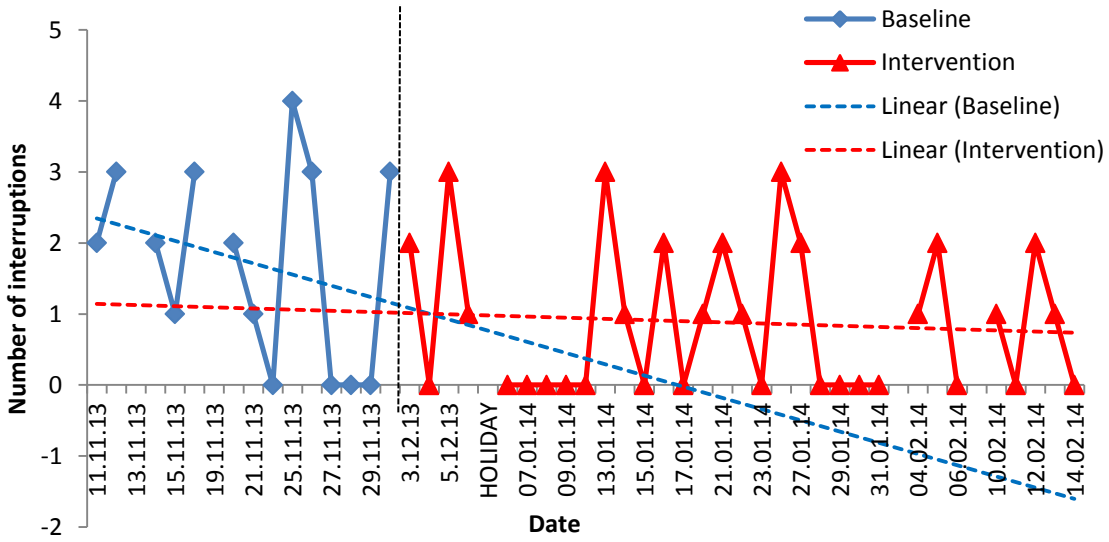


Figure 5-8: A graph showing the number of interruptions of the teacher made by S1a each day, including trend lines.

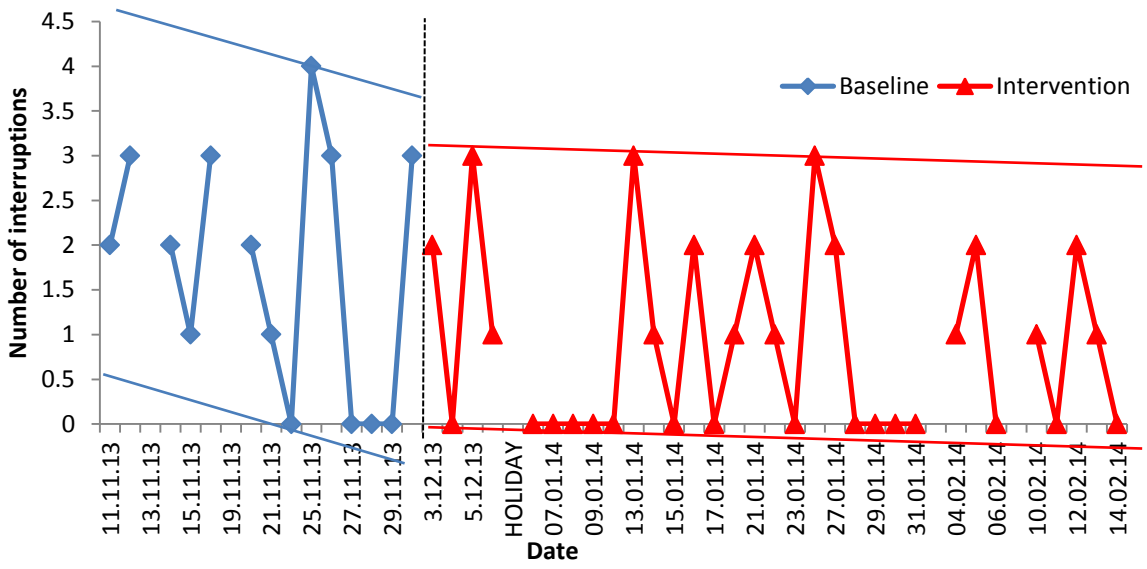


Figure 5-9: A graph showing the number of interruptions of the teacher made by S1a each day, including variability lines.

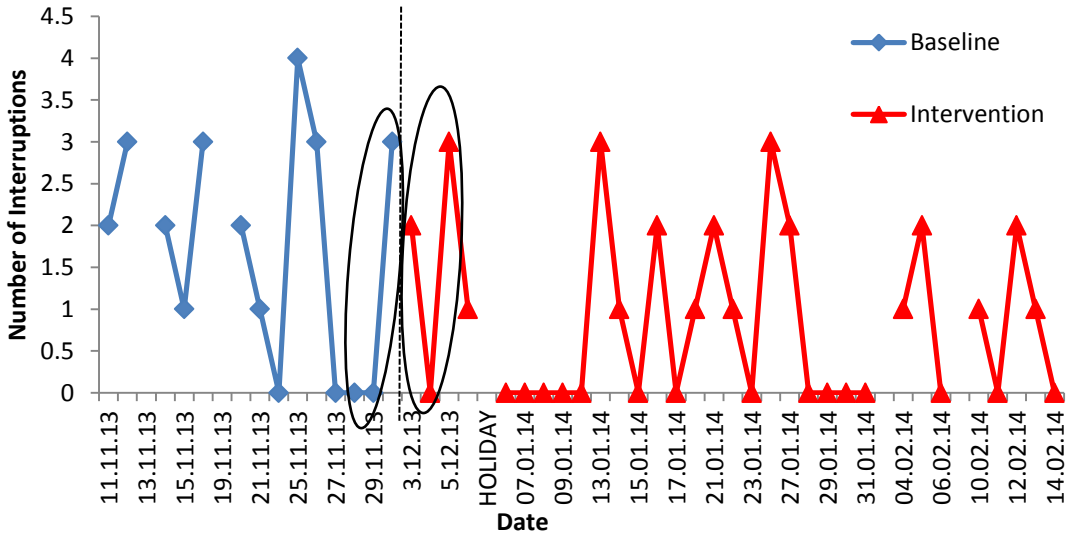


Figure 5-10: A graph showing the number of interruptions of the teacher made by S1a each day, showing immediacy of effect.

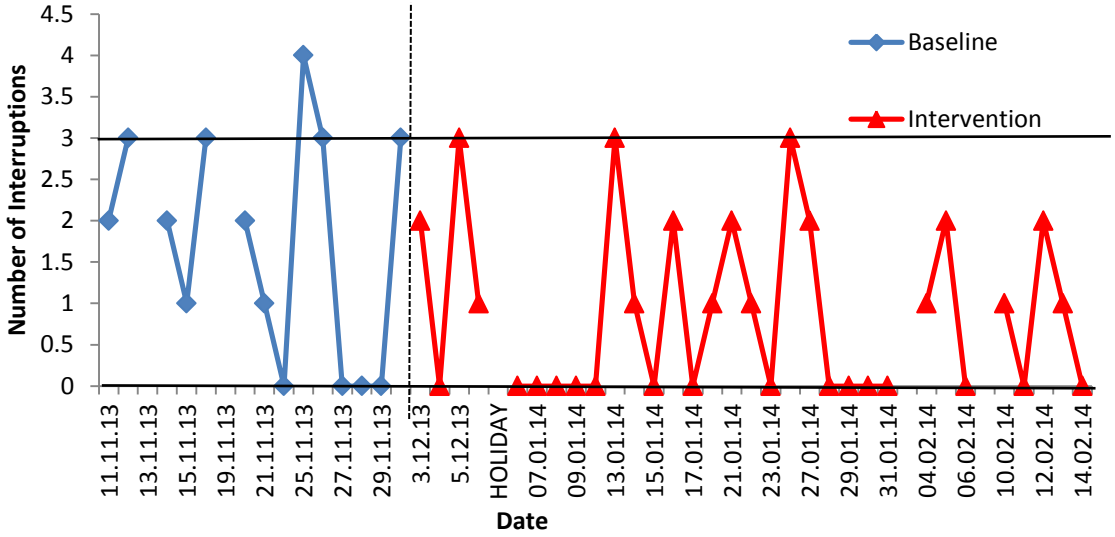


Figure 5-11: A graph showing the number of interruptions of the teacher made by S1a each day, including lines to show overlap.

Table 5-30: A table showing the standard deviation and mean for number of interruptions made by S1a per day.

	Baseline	**Intervention**	**Difference**
Mean	1.71	0.88	0.83
Standard Deviation	1.33	1.07	0.26

Table 5-31: A table outlining brief descriptions of the outcome of the visual analysis of teacher interruptions made by S1a.

Feature	**Visual Analysis**
Level	Figure 5-7 shows a decrease in the mean level of interruptions (reducing by 0.83 (Table 5-30)) after implementation of the intervention.
Trend	Figure 5-8 shows a sharp decrease in behaviour during the baseline phase. The intervention trend has a very small decline, it begins slightly lower than the baseline trend.
Variability	The variability lines (Figure 5-9) show a narrower variance for the intervention phase than the baseline phase.
Immediacy of Effect	The immediacy of effect of intervention does not show a clear pattern (but suggests an increase of the behaviour) in behaviour change compared to the final 3 points in the baseline (Figure 5-10).
Overlap	The lowest point in both phases is the same (0) (Figure 5-11). The baseline has a higher data point in 1 of the 14 data points (7%).

5.2.4.3 S1a – Behaviour (Complaints)

Data is presented graphically with key numerical information presented in Table 5-32.

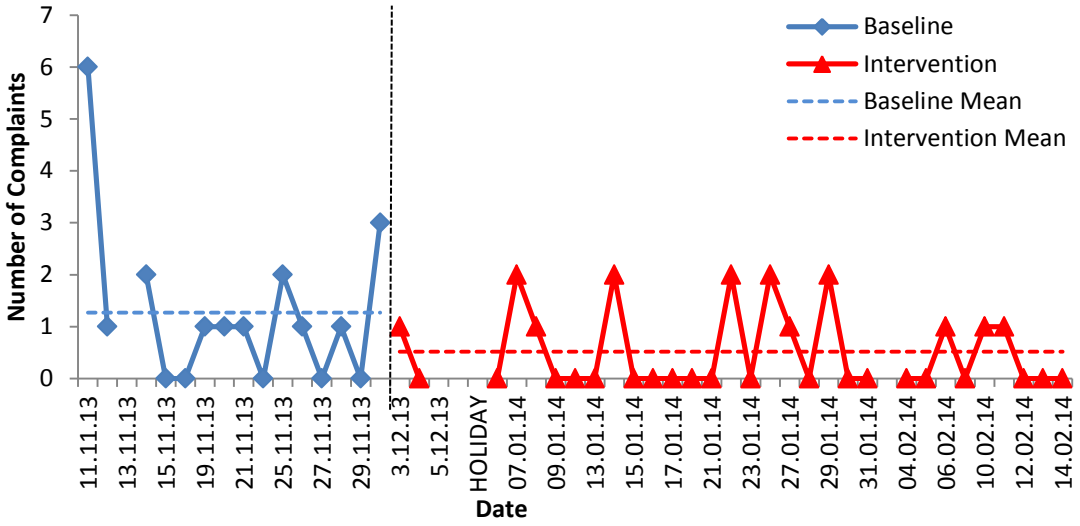


Figure 5-12: A graph showing the number of complaints made daily by S1a, including level lines.

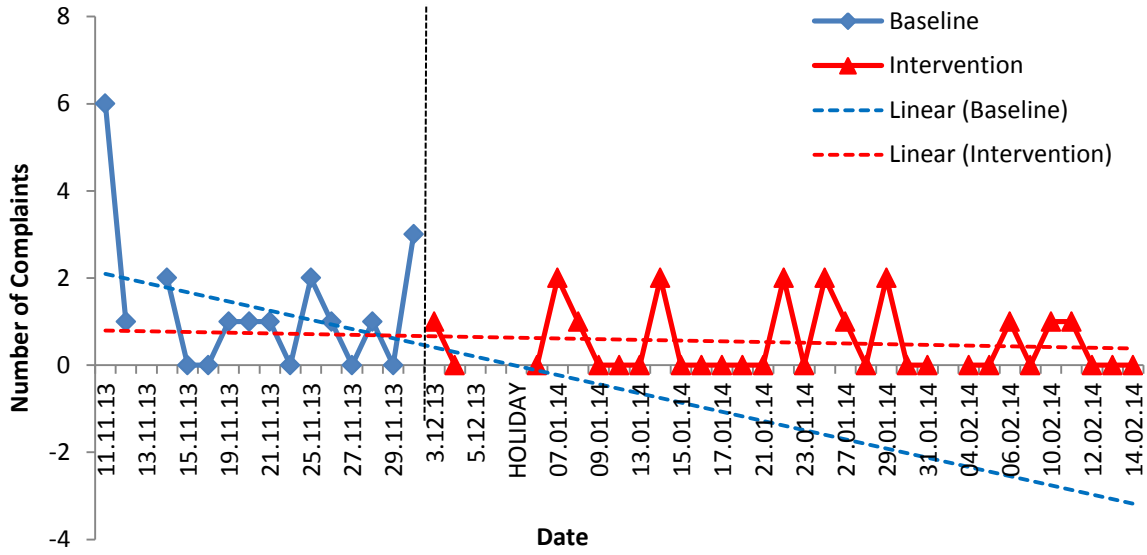


Figure 5-13: A graph showing the number of complaints made daily by S1a, including trend lines.

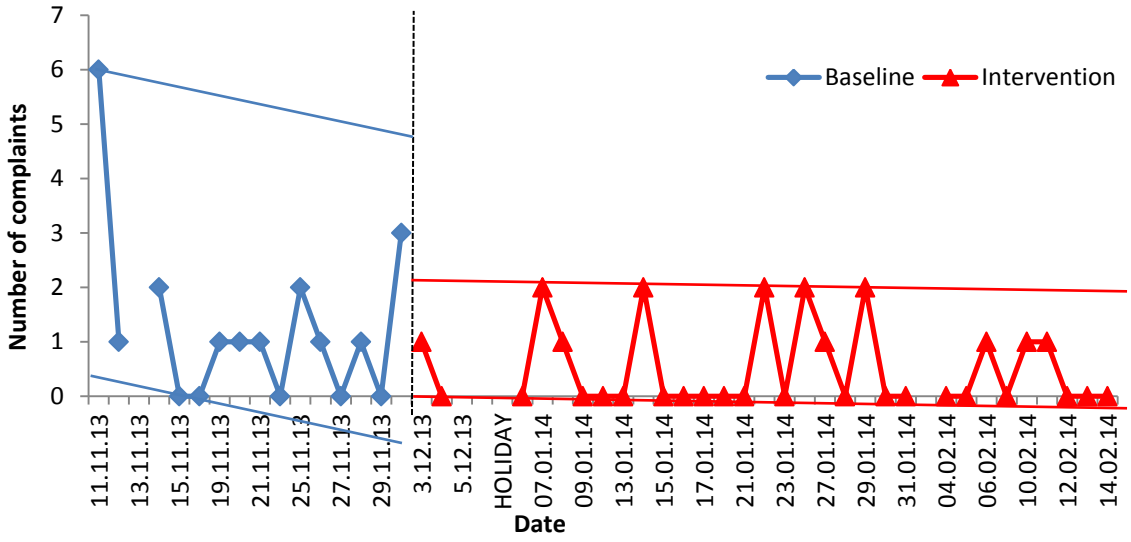


Figure 5-14: A graph showing the number of complaints made daily by S1a, including variability lines.

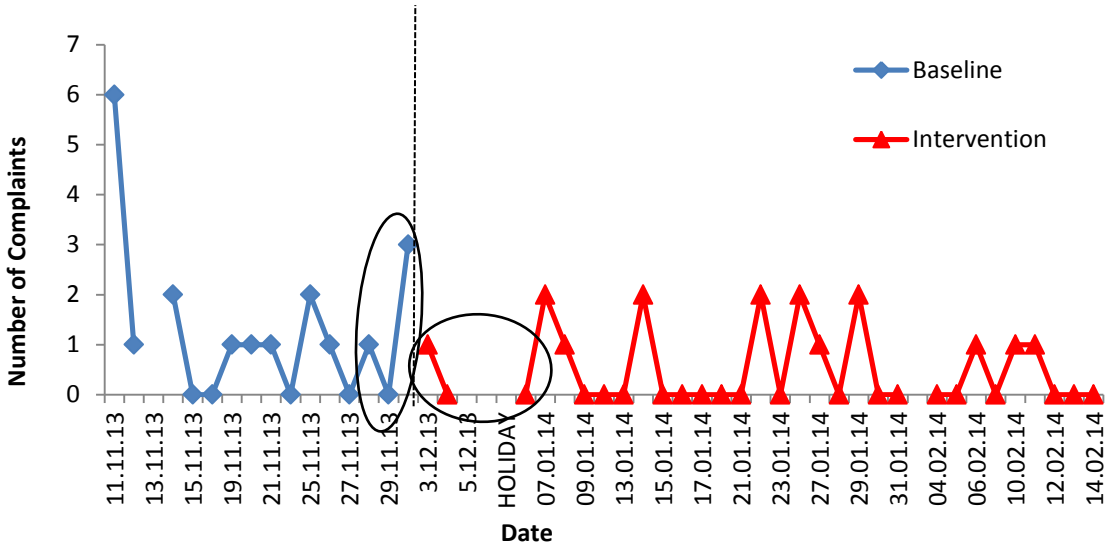


Figure 5-15: A graph showing the number of complaints made daily by S1a, including highlighted immediacy of effect.

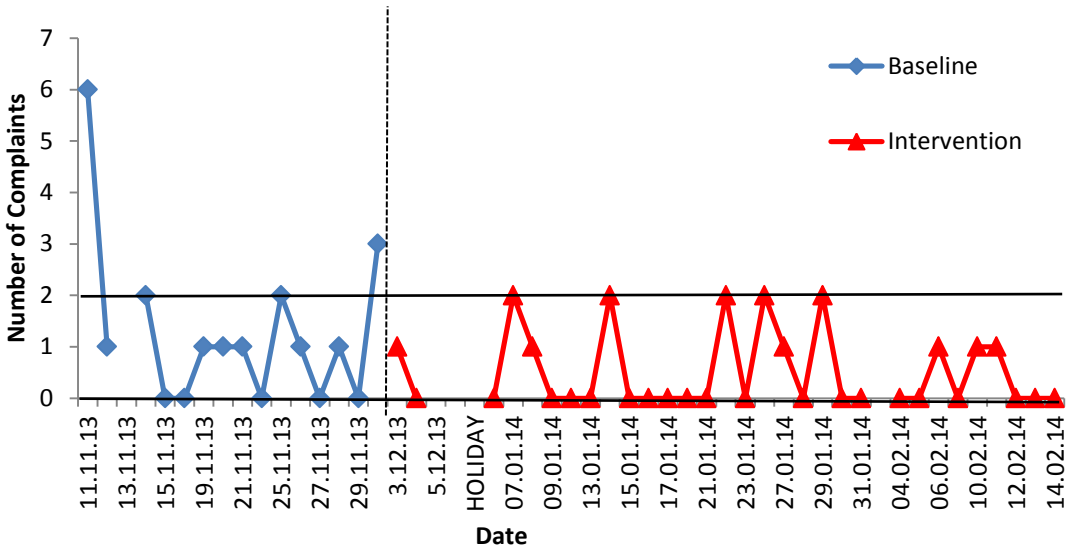


Figure 5-16: A graph showing the number of complaints made daily by S1a, including overlap.

Table 5-32: A table showing the standard deviation and mean for number of complaints made by S1a per day.

	Baseline	**Intervention**	**Difference**
Mean	1.27	0.52	-0.75
Standard Deviation	1.53	0.80	-0.73

Table 5-33: A table outlining brief descriptions of the outcome of the visual analysis of daily complaints made by S1a.

Feature	**Visual Analysis**
Level	Figure 5-12 shows a decrease in mean level during the intervention phase compared to the baseline. This is a difference of 0.75 (Table 5-32).
Trend	The baseline trend line shows a sharp decline. This is due to the first data point being much higher than the other data points. The baseline trend line also shows a decrease in behaviours but with only a slight gradient. The intervention phase trend line begins much lower than the baseline trend line (Figure 5-13).
Variability	Figure 5-14 shows a larger variability for the baseline phase than the intervention phase. This is as large as it is due to an outlier in the baseline phase. Without this outlier the variability in the baseline would still be larger, although the difference between the phases would be smaller.
Immediacy of Effect	Immediacy of effect suggests a decline in behaviour in the intervention phase compared to the baseline phase (Figure 5-15).
Overlap	The lowest of data points in both phases is 0. The data in the baseline phase is higher than the intervention phase for 2 (0.13%) of its data points.

5.2.4.4 S1a – Behaviour 4 (Refusal)

Data is presented graphically with key numerical information presented in Table 5-34.

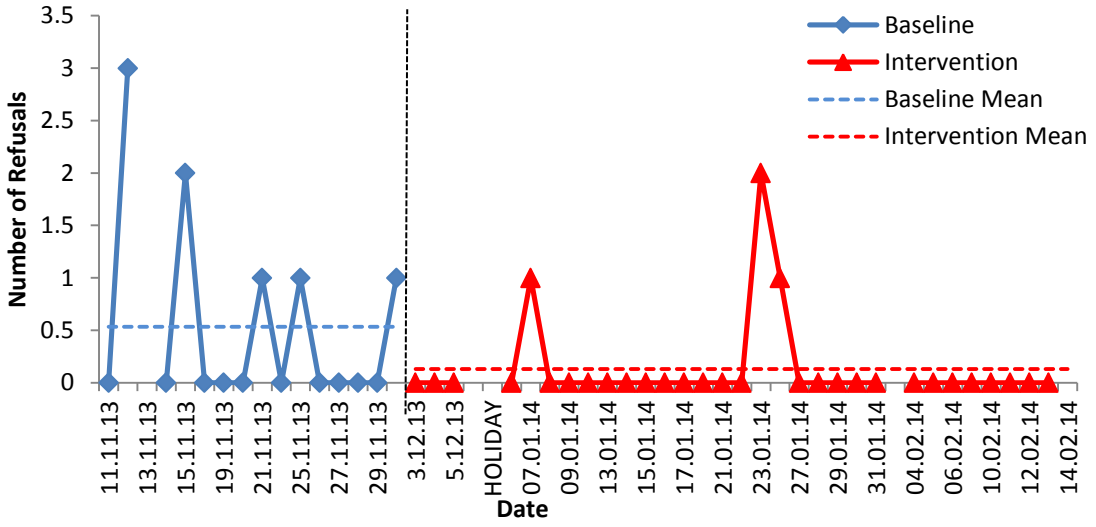


Figure 5-17: A graph showing the number of refusals made daily by S1a, including mean lines.

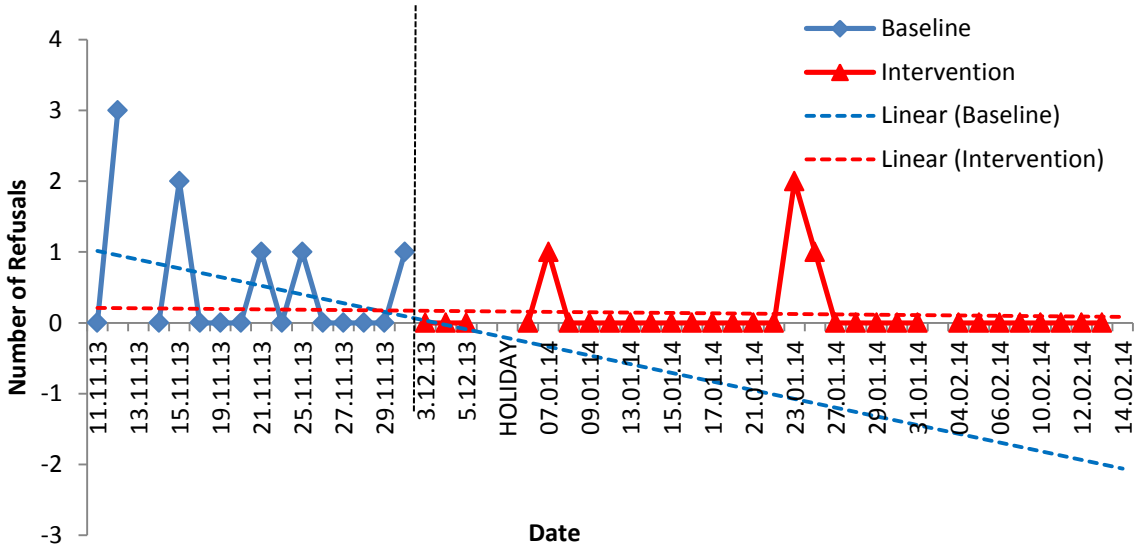


Figure 5-18: A graph showing the number of refusals made daily by S1a, including trend lines.

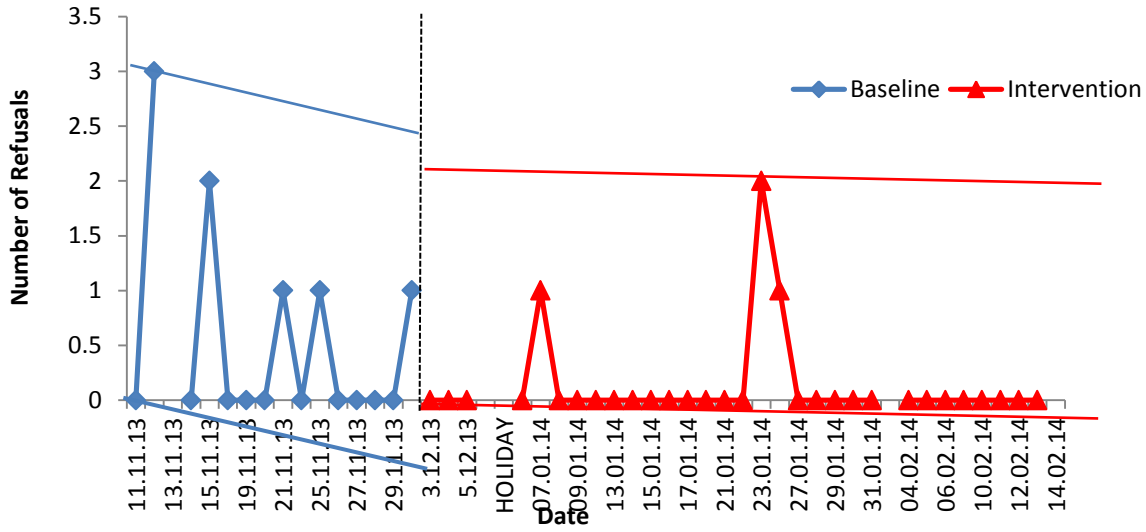


Figure 5-19: A graph showing the number of refusals made daily by S1a, including variability lines

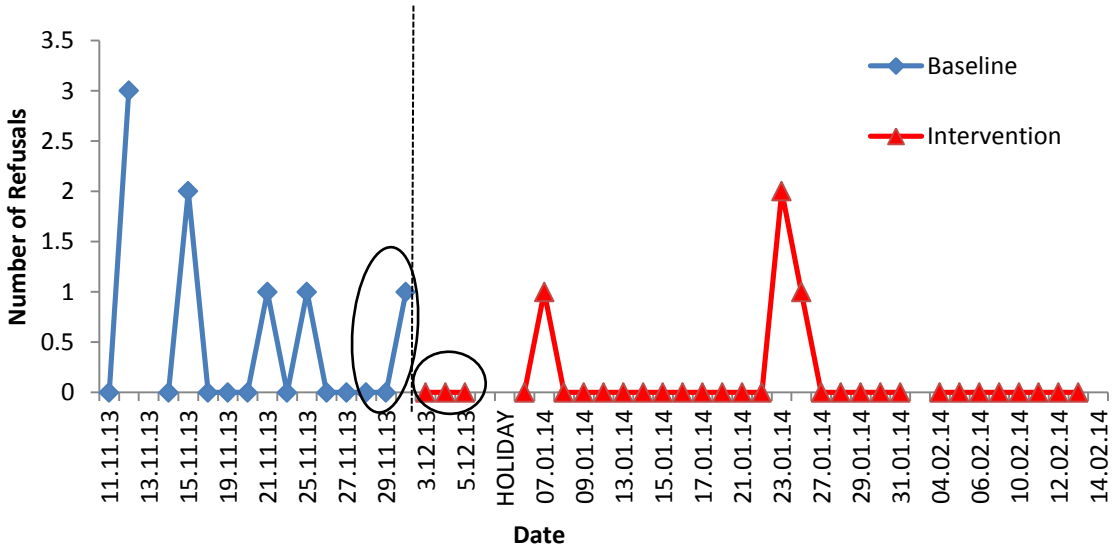


Figure 5-20: A graph showing the number of refusals made daily by S1a, including highlighted immediacy of effect.

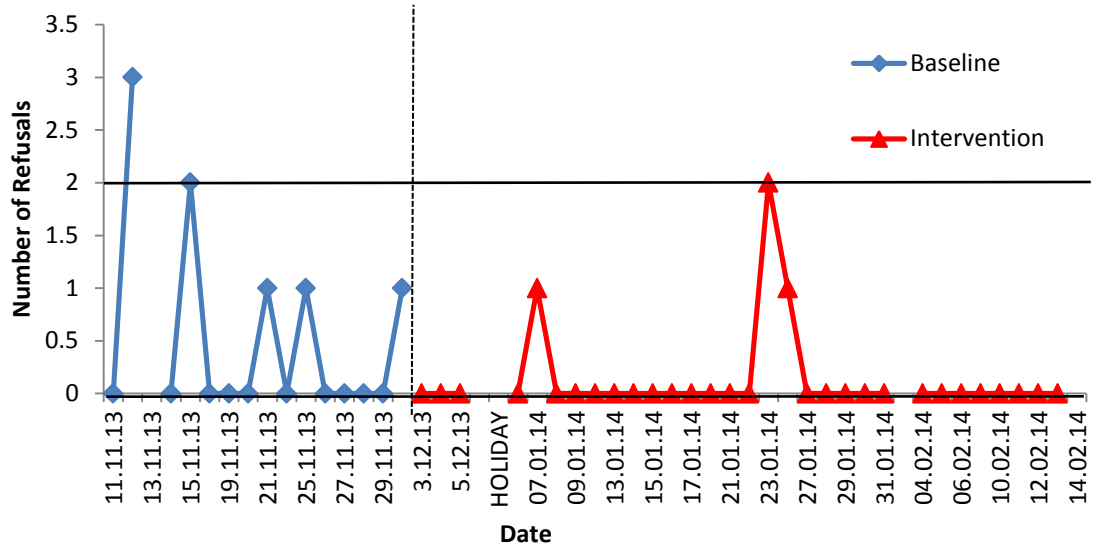


Figure 5-21: A graph showing the number of refusals made daily by S1a, including overlap lines.

Table 5-34: A table showing the standard deviation and mean for number of refusals made by S1a per day.

	Baseline	**Intervention**	**Difference**
Mean	0.53	0.13	0.40
Standard Deviation	0.88	0.46	0.42

Table 5-35: A table outlining brief descriptions of the outcome of the visual analysis of daily refusals made by S1a.

Feature	**Visual Analysis**
Level	The level is notably lower in the intervention phase compared to the baseline phase (Figure 5-17). There is a difference of 0.4 refusals per day (Table 5-34).
Trend	The trend line in the baseline phase shows a steep decline in behaviours (Figure 5-18). The intervention trend line begins much lower and has a very shallow decrease due to the large number of zeros.
Variability	The variability is larger in the baseline phase than the intervention phase (Figure 5-19).
Immediacy of Effect	The immediacy of effect data points in the intervention are all '0'. The final 3 stages of the baseline also include '0' but this spikes to '1' for the final point.
Overlap	Both phases reach the lowest data point (0). The baseline phase has 1 data point higher than occurs in the intervention phase.

5.2.5 S1a – Summary of Analysis

The data is a little unclear for behaviour 1 (inappropriate response to not being chosen). This lack of clarity is very much impacted by the outlier in the intervention phase. With this outlier included in the analysis the data shows a decline in the behaviour for level only, suggesting no impact of intervention. However, excluding the outlier (see 6.2 for possible explanations), the results suggest a decrease in behaviour and increased stability according to variance, level, and overlapping data points after implementation of the intervention. This results in 3 effects, 1 non-effect (immediacy of effect), and an unclear result (trend). With 3 effects and 1 zero effect the intervention could be said to have "moderate effect" (Kratochwill et al, 2011).

Behaviour 2 (interrupting teacher) shows a decrease in presentation after implementation in the mean. There is less variability shown in the intervention phase and the overlap also suggests that behaviour has decreased. The immediacy of effects has no clear pattern. The trend suggests although the intervention phase is lower overall, the gradient of decline is higher in the baseline phase. These results are likely to have been affected by the instability of baseline. This analysis shows small to no effect of the intervention on the behaviour.

Behaviour 3 (complaints made) suggests that the behaviour decreases after intervention implementation according to level, variability, immediacy of effect, and overlap. The trend lines are unclear. This result would suggest a "strong effect" of intervention. However, the researcher feels that this should be interpreted with caution due to the impact of the unstable baseline on analysis and therefore concludes a moderate effect of intervention on the number of complaints made.

Behaviour 4 (refusals) data shows a decline in behaviours after the intervention implementation in level, variability, and immediacy of effect. Trend and overlap data do not create a clear analysis of the data. Consequently this data could support a hypothesis of small to moderate impact of the intervention.

The data suggests a small to moderate effect of the intervention on 3 of the 4 behaviours recorded for S1a (an effect is also shown for behaviour 1 if the outlier is removed from analysis). The strongest impact appears to be on S1a's inappropriate reaction to being chosen and complaints made to the teacher. The data for interruptions and refusals present a less clear picture. All analyses are impacted by the instability of the baselines.

5.2.6 Case 2 (S1b)

Name: S1b

Age: 8 years at start of intervention

Year Group: Year 3

Behaviours Recorded:

Behaviour 1: Inappropriate crying (number and length)

Definition: S1b cries and complains loudly in a manner which is louder and longer than expected from peers his age in response to similar events (see Appendix 7 for full definition).

Behaviour 2: Inappropriate noises during work-time

Definition: S1b makes loud vocalisations / noises when he is supposed to be listening / quiet (see Appendix 7 for full definition).

5.2.6.1 S1b – Behaviour 1 (Crying – number of incidents)

Data is presented graphically with key numerical information presented in Table 5-36.

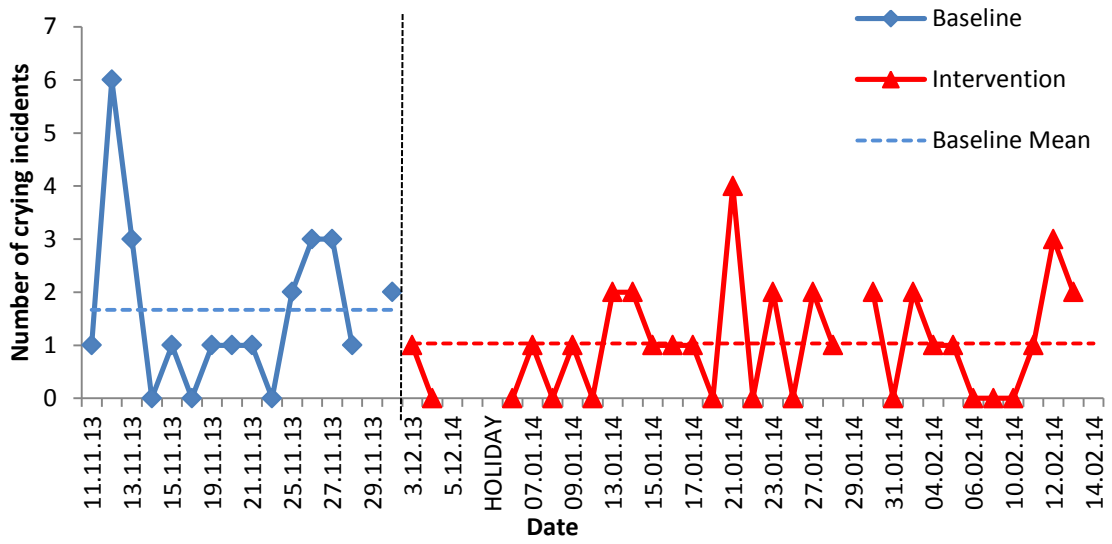


Figure 5-22: A graph showing the number of incidents of inappropriate crying for S1b each day, including level lines.

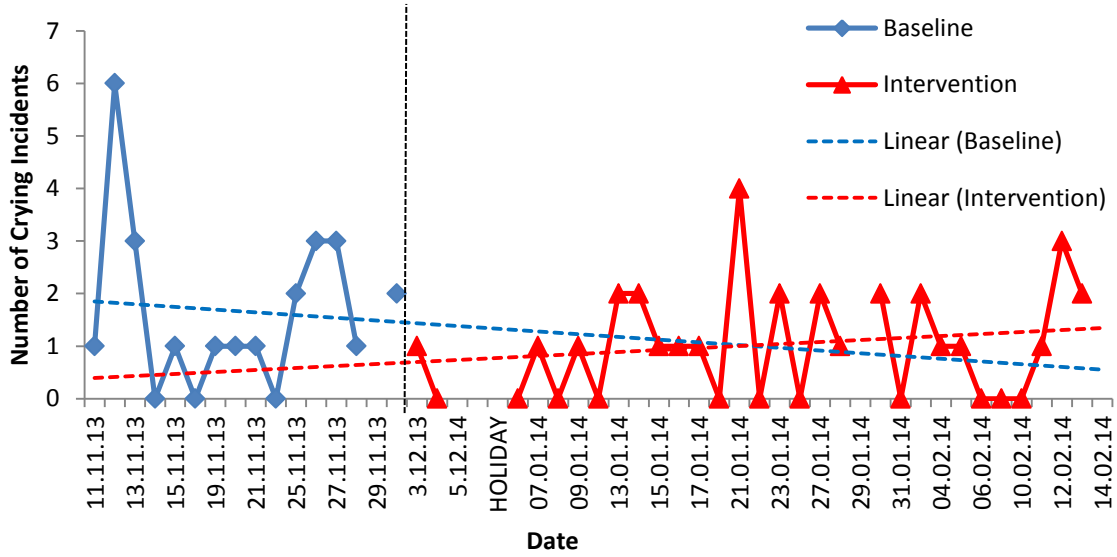


Figure 5-23: A graph showing the number of incidents of inappropriate crying for S1b each day, including trend lines.

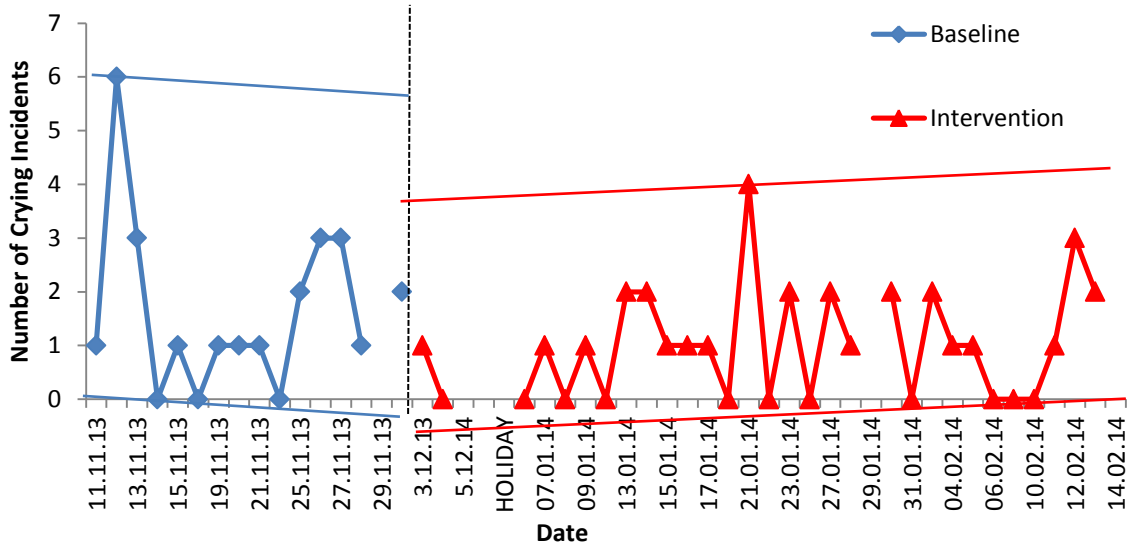


Figure 5-24: A graph showing the number of incidents of inappropriate crying for S1b each day, including variability lines.

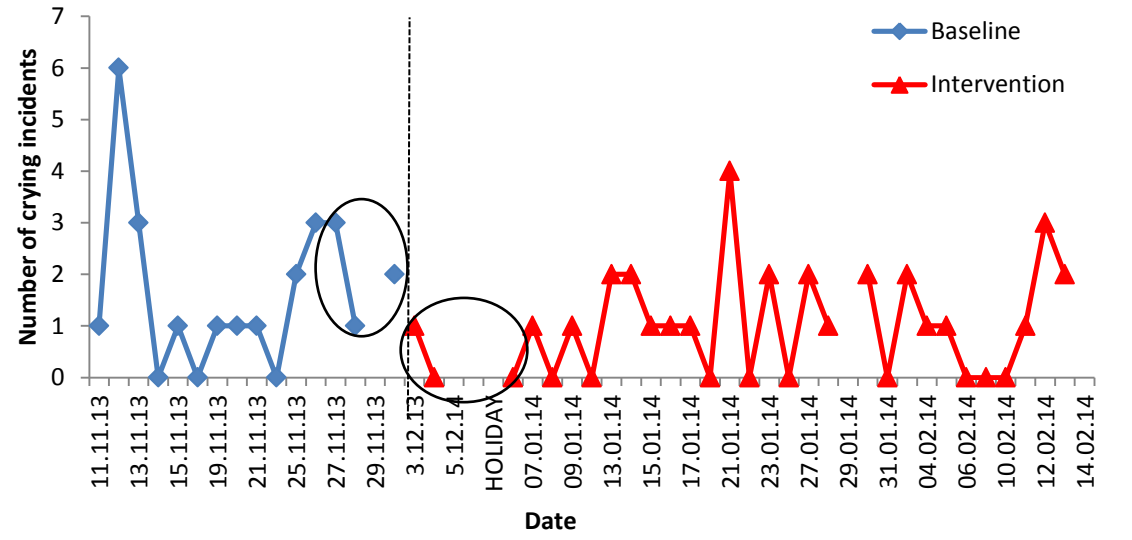


Figure 5-25: A graph showing the number of incidents of inappropriate crying for S1b each day, including immediacy of effect.

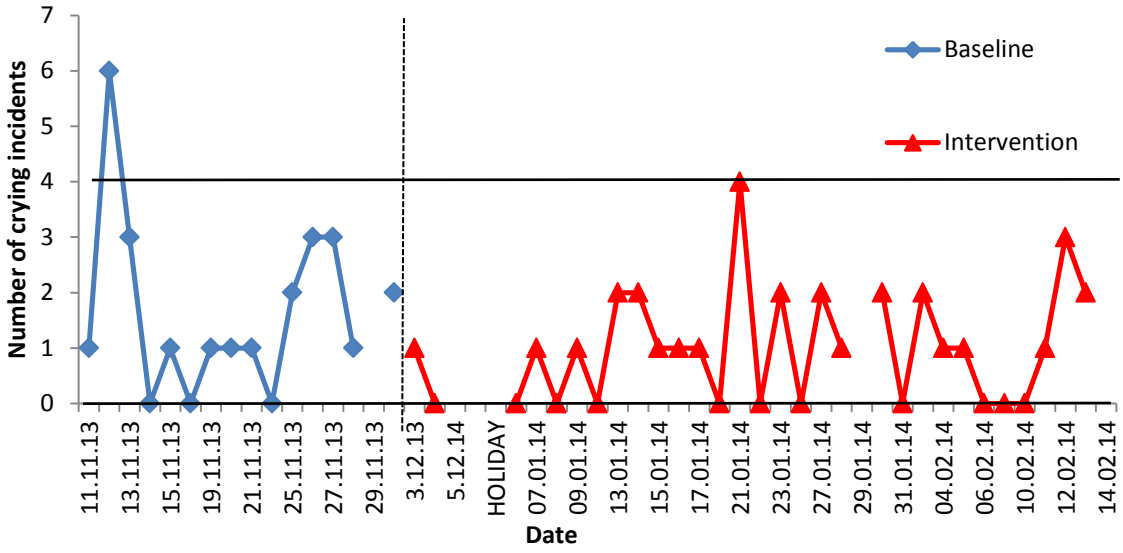


Figure 5-26: A graph showing the number of incidents of inappropriate crying for S1b each day, with overlapping data lines.

Table 5-36: A table showing the means and standard deviations for the number of times S1a cries inappropriately each day.

	Baseline	**Intervention**	**Difference**
Mean	1.67	1.03	0.64
Standard Deviation	1.53	1.02	0.51

Table 5-37: A table outlining brief descriptions of the outcome of the visual analysis of number of inappropriate crying incidents for S1b.

Feature	**Visual Analysis**
Level	The level is lower for the intervention phase than the baseline phase (Figure 5-22). The level is 0.64 lower in the intervention phase (Table 5-36).
Trend	The trend shows the number of incidents of inappropriate crying to increase during the intervention. The number of incidents appeared to be decreasing during the baseline (Figure 5-23).
Variability	The variability is shown to be much greater in the baseline phase than the intervention phase (Figure 5-24).
Immediacy of Effect	The first 3 data points in the intervention phase are all lower than the final 3 of the baseline (Figure 5-25). There is also an immediate drop from the first data point to the second and the third (0) in the intervention.
Overlap	The lowest data points are shared between the baseline and the intervention phase. Only 1 data point out of 15 (33.33%) is higher in the baseline phase than any point in the intervention phase.

5.2.6.2 S1b – Behaviour 1 (Crying – time)

Data is presented graphically with key numerical information presented in Table 5-38

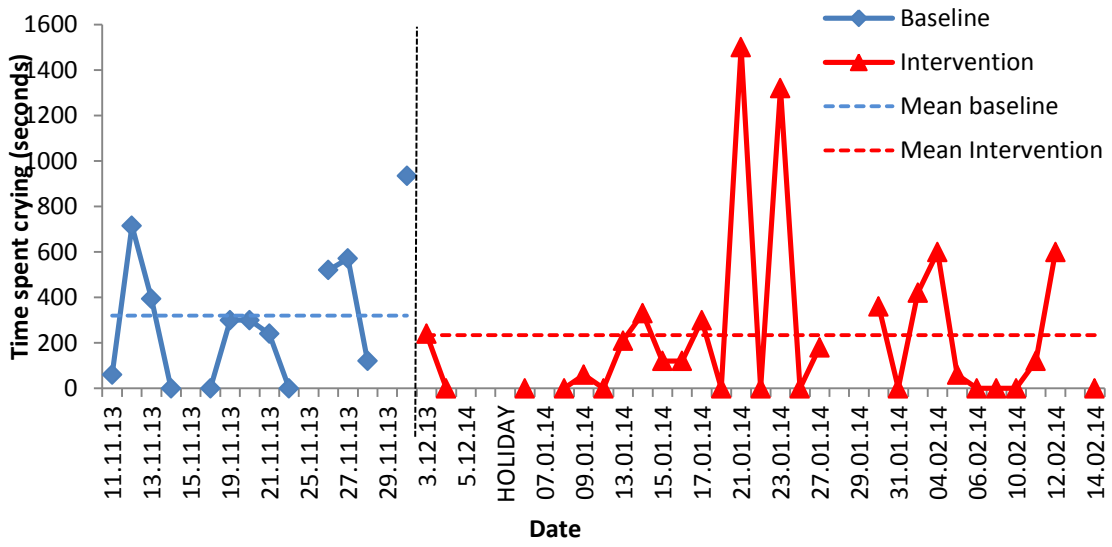


Figure 5-27: A graph showing the amount of time spent crying inappropriately by S1b each day, including level lines.

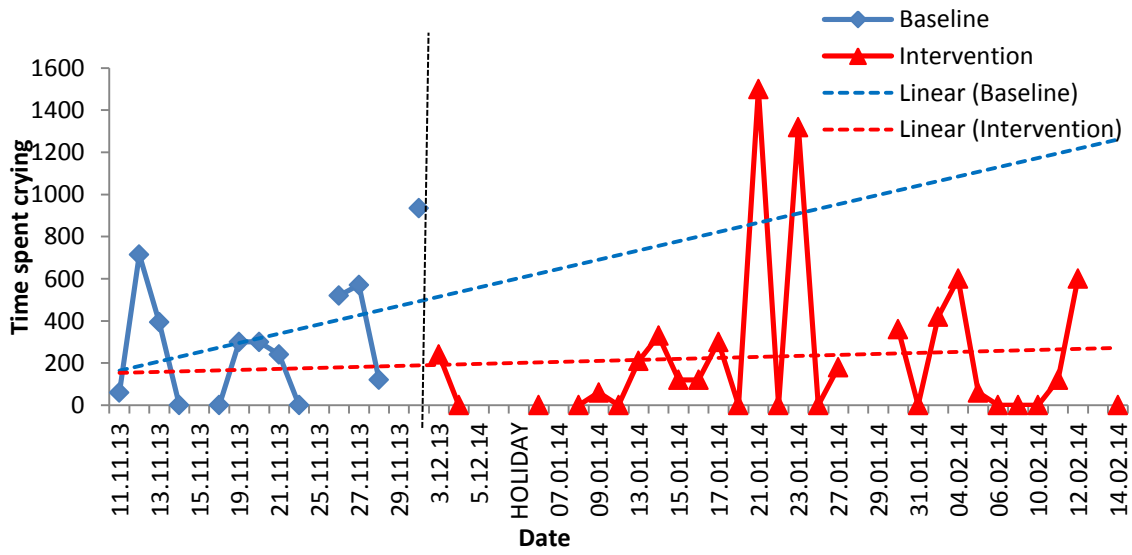


Figure 5-28: A graph showing the amount of time spent crying inappropriately by S1b each day, including trend lines.

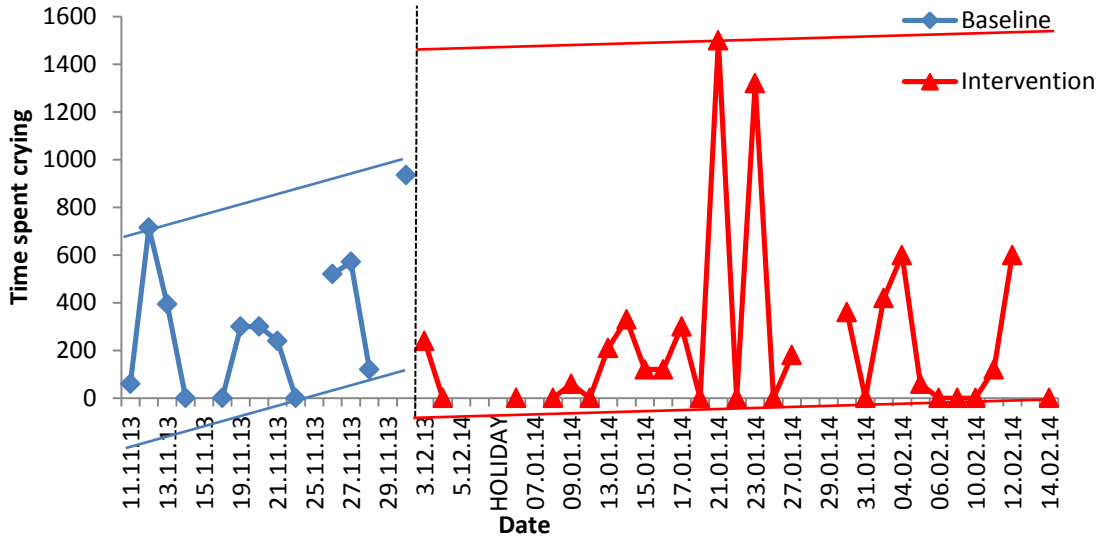


Figure 5-29: A graph showing the amount of time spent crying inappropriately by S1b each day, including variability lines.

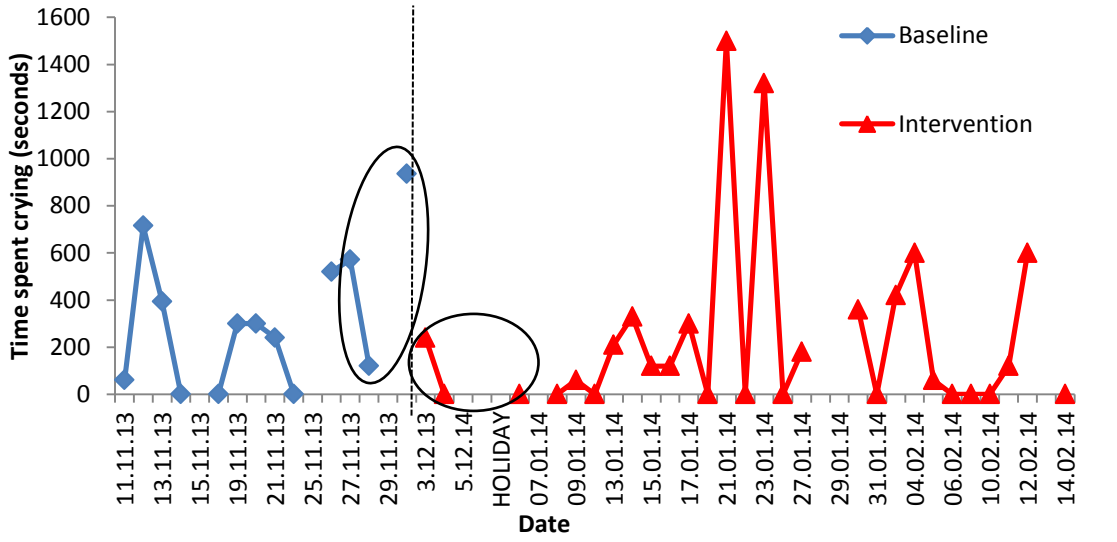


Figure 5-30: A graph showing the amount of time spent crying inappropriately by S1b each day, with immediacy of effect highlighted.

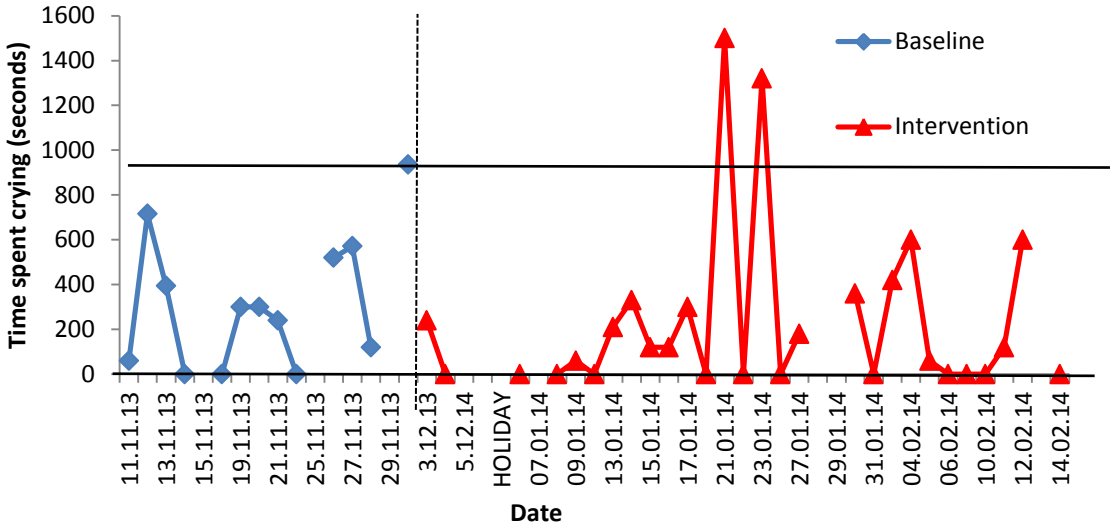


Figure 5-31: A graph showing the amount of time spent crying inappropriately by S1b each day, with marked overlap.

Table 5-38: A table showing the means and standard deviations for the amount of time S1b spends crying inappropriately each day.

	Baseline	**Intervention**	**Difference**
Mean	319.62	233.57	86.05
Standard Deviation	286.19	371.98	-85.79

Table 5-39: A table outlining brief descriptions of the outcome of the visual analysis of amount of time S1b spends crying inappropriately each day.

Feature	**Visual Analysis**
Level	The level is lower during the baseline than the intervention (Figure 5-27). The level is 86.05 seconds lower (Table 5-38).
Trend	Both trend lines suggest the amount of time S1b spent crying each day was increasing (Figure 5-28). The trend line for the intervention phase has a markedly shallower gradient.
Variability	Figure 5-29 shows a much greater variability in the intervention phase than the baseline phase.
Immediacy of Effect	Figure 5-30 shows the first 3 data points of the intervention phase to be much lower than the final points of the baseline phase. The first 3 points in the intervention phase decrease over time.
Overlap	There are 2 very high outliers (7%) in the intervention phase which are higher than any points in the baseline phase. Both phases go down to 0 (Figure 5-31).

5.2.6.3 S1b – Behaviour 2 (Inappropriate noises)

Data is presented graphically with key numerical information presented in Table 5-40.

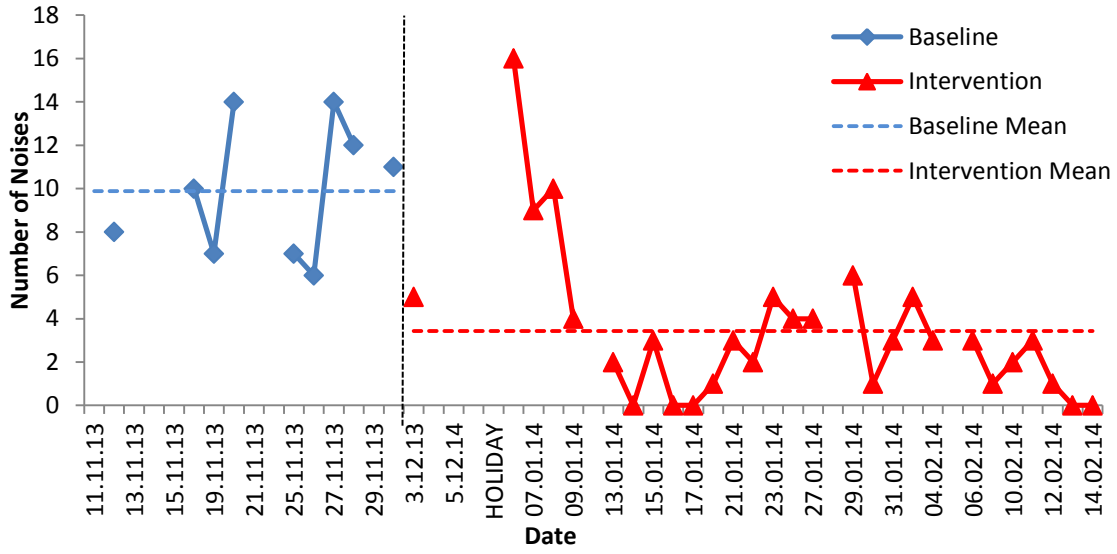


Figure 5-32: A graph showing the number of inappropriate noises made by S1b in class each day, including mean lines.

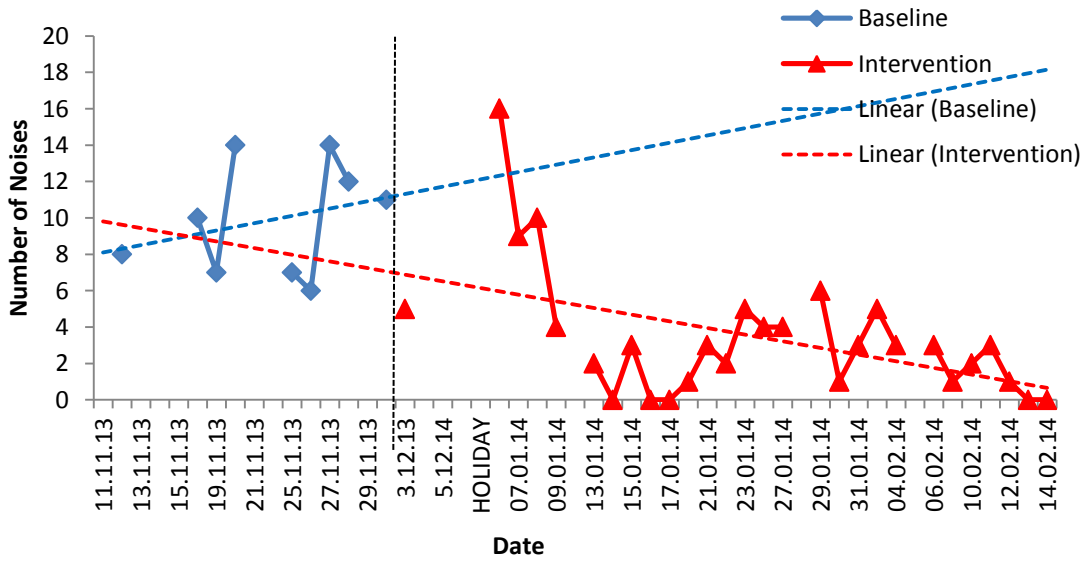


Figure 5-33: A graph showing the number of inappropriate noises made by S1b in class each day, including trend lines.

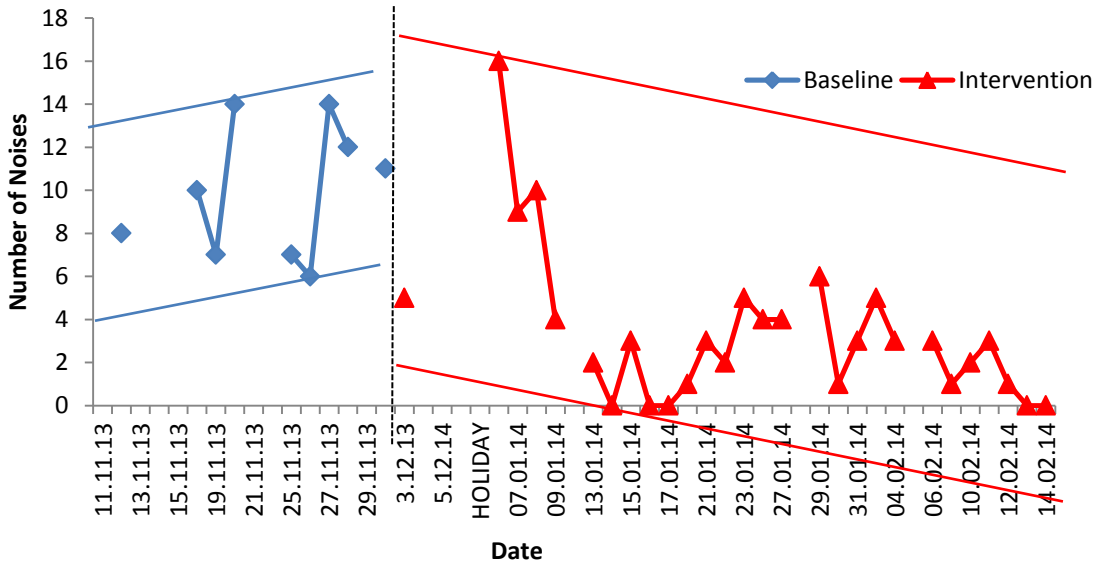


Figure 5-34: A graph showing the number of inappropriate noises made by S1b in class each day, including variability lines.

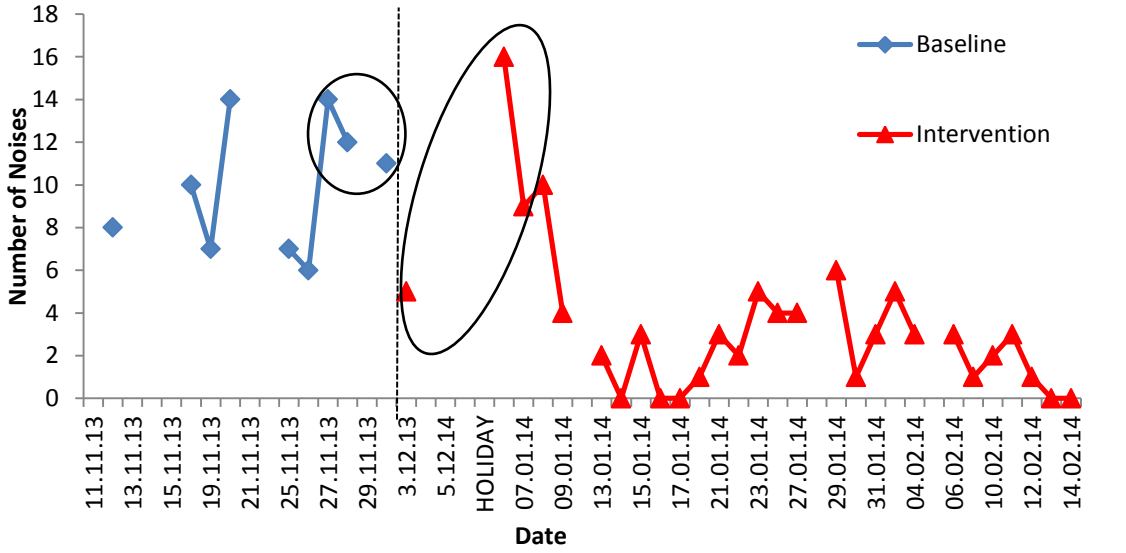


Figure 5-35: A graph showing the number of inappropriate noises made by S1b in class each day with immediacy of effect highlighted.

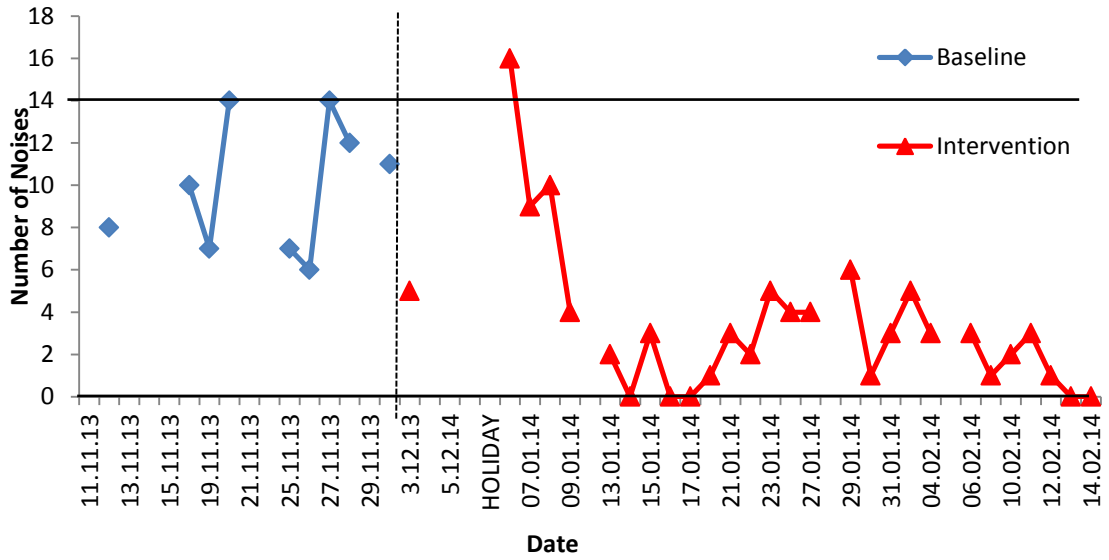


Figure 5-36: A graph showing the number of inappropriate noises made by S1b in class each day with percentage of data overlap.

Table 5-40: A table showing the means and standard deviations for the number of times S1b makes inappropriate noises each day.

	Baseline	**Intervention**	**Difference**
Mean	9.89	3.43	6.46
Standard Deviation	2.88	3.47	-0.59

Table 5-41: A table outlining brief descriptions of the outcome of the visual analysis of amount of inappropriate noises made by S1b.

Feature	**Visual Analysis**
Level	The level is noticeably lower during the intervention phase than the baseline phase (Figure 5-32). Table 5-40 shows this difference to be 6.47 less noises per day in the intervention phase.
Trend	Table 5-33 shows that the behaviour is increasing sharply over the baseline period. The intervention trend line shows a steep decline in inappropriate noises.
Variability	The variability of the behaviour is greater in the intervention phase than the baseline phase (Table 5-34).
Immediacy of Effect	The immediacy of effect is unclear (Figure 5-35). The data shows an immediate decline from baseline to intervention, followed by a steep incline and then a decrease to the third data point, which continues.
Overlap	The overlap data shows 1 of the intervention points to be higher than any point in the baseline (4%). The intervention data is lower than the intervention data for 25 of the 28 data points (89%).

5.2.7 S1b - Summary of Analysis

Behaviour 1 (inappropriate crying) was measured in number of incidents as well as amount of time spent crying (seconds). The data for the number of incidents was seen to decrease for 3 of the analysis features (level, variability, and immediacy of effect). The trend does not support the hypothesis as the number of incidents decreased during the baseline and increased during the intervention. The overlap data suggests a decrease of incidents but this relies on only 1 data point so is a little unclear. The amount of time spent crying decreased according to level, trend, and immediacy of effect. However, the variability was greater in the intervention phase and 7% of data points were higher during the intervention phase compared to the baseline phase. Therefore it can be suggested the intervention had a moderate effect on number of incidents of crying with an unclear or no effect on the amount of time spent crying.

Visual analysis of the features for behaviour 2 (inappropriate noises) suggests a moderate to strong impact of the intervention on the behaviour in the desired direction. The only feature which does not suggest a decrease in behaviours is immediacy of effect and this is unclear as oppose to showing "no-effect" (see Table 5-41). The data therefore suggests a strong to moderate effect of intervention on decreasing inappropriate noises.

5.2.8 Case 3 (S1c)

Name: S1c

Age: 7 years and 4 months at the start of intervention

Year Group: Year 1

Behaviours Recorded:

Behaviour: Refusal to comply with an instruction given by a staff member (number of refusals per day and seconds spent refusing)

Definition: S1c refuses to comply with an instruction from a staff member after 5 seconds (see Appendix 7 for full definition).

5.2.8.1 S1c – Behaviour 1 (Number of Refusals)

Data is presented graphically with key numerical information in Table 5-42.

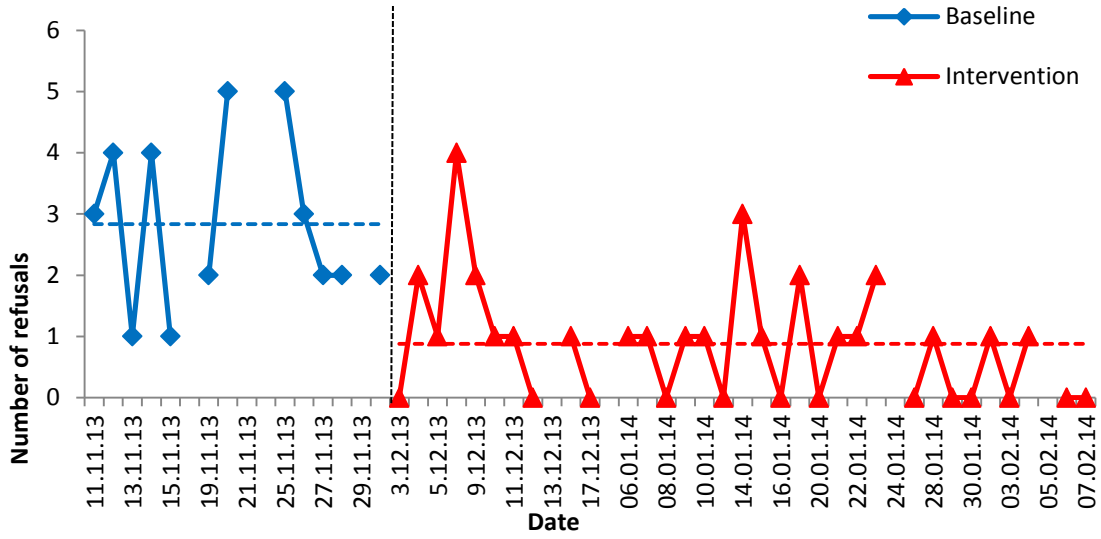


Figure 5-37: A graph showing the number of refusals per day and mean number of refusals for participant S1c across Phase A and B.

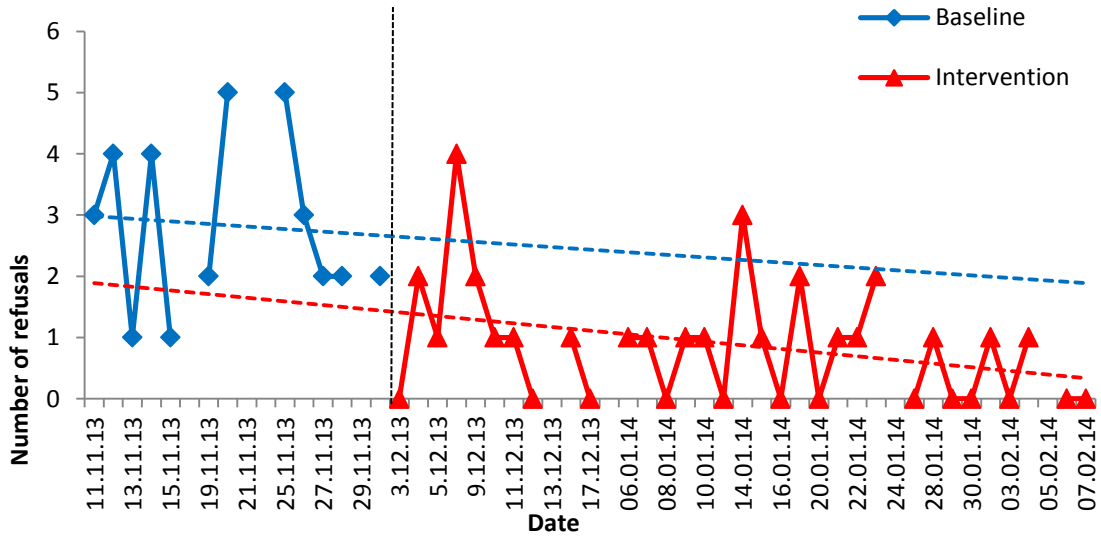


Figure 5-38: A graph showing the number of refusals per day and trend lines for each phase for participant S1c across Phase A and B.

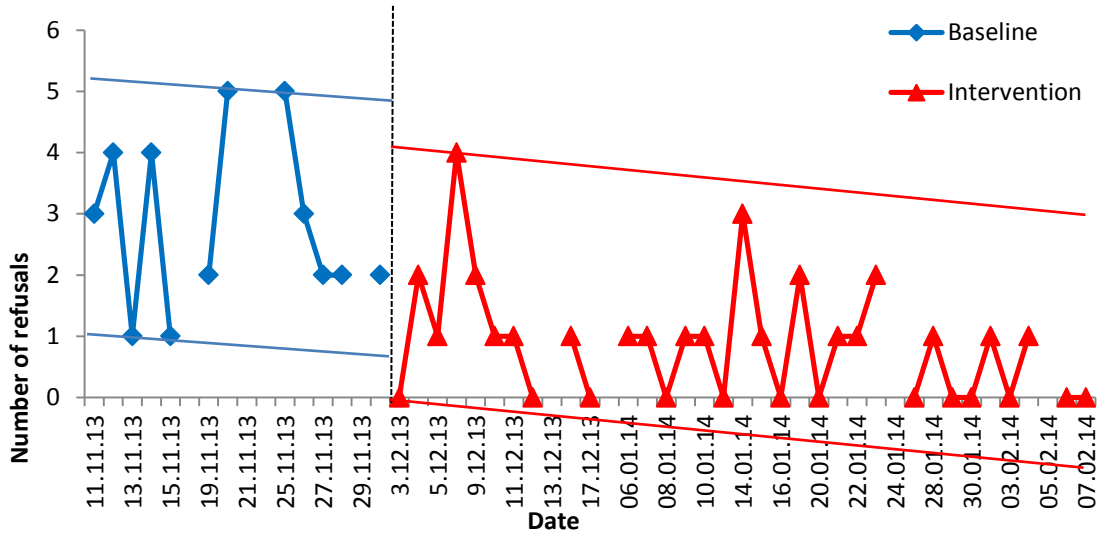


Figure 5-39: A graph showing the number of refusals per day and variability lines for each phase for participant S1c across Phase A and B.

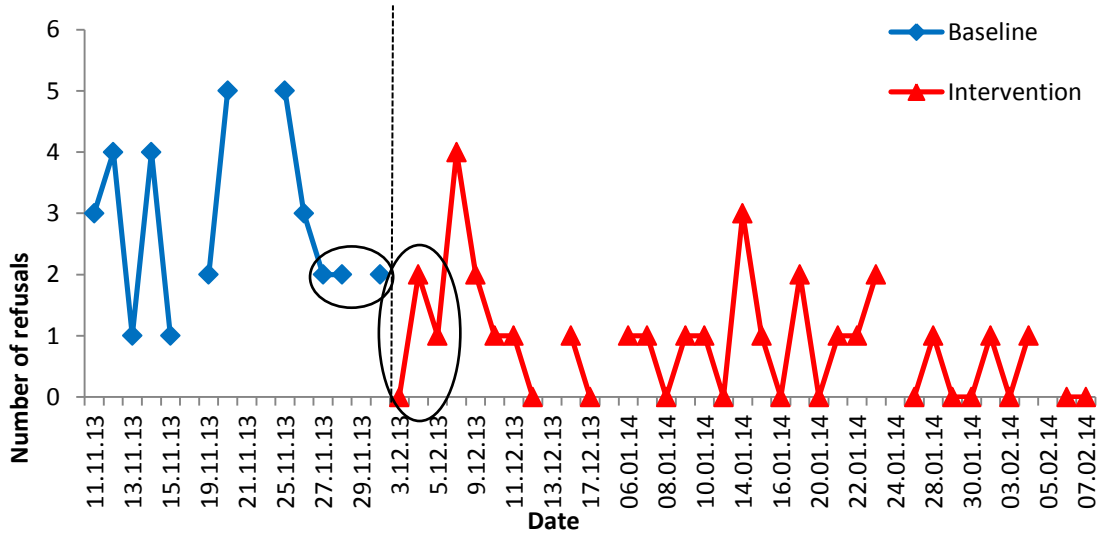


Figure 5-40: A graph showing the number of refusals per day for participant S1c across Phase A and B and immediacy of effect.

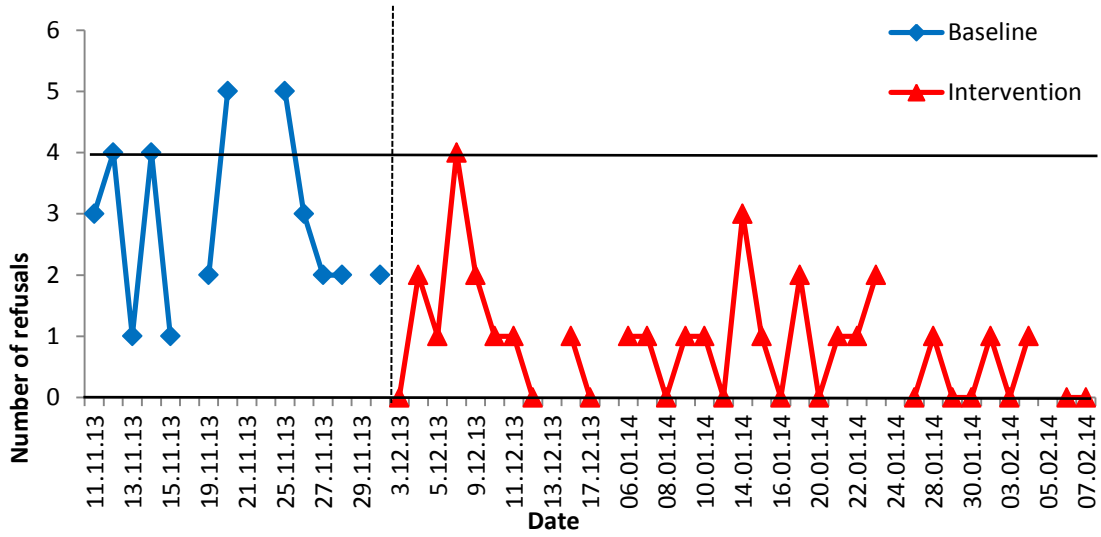


Figure 5-41: A graph showing the number of refusals per day and overlap for each phase for participant S1c across Phase A and B.

Table 5-42: A table showing key numerical information for S1c (number of refusals per day).

	Baseline	**Intervention**	**Difference**
Mean	2.83	0.88	-1.95
Standard Deviation	1.34	0.95	-0.39

Table 5-43: A table outlining brief descriptions of the outcome of the visual analysis of number of refusals for S1c.

Feature	**Visual Analysis**
Level	Figure 5-37 shows a notable decrease in the mean level (reducing by 1.95 (Table 5-42)).
Trend	The trend lines (Figure 5-38) show that the refusals per day were reducing in both baseline and intervention. However, the gradient in the intervention phase is steeper than in the baseline.
Variability	The variability lines (Figure 5-39) show that the variance in the intervention is smaller in the intervention phase than the baseline phase.
Immediacy of Effect	Figure 5-40 shows that the number of refusals is lower in the intervention for point 1 and 3. However, the second point is as high as the 3 points in the intervention phase.
Overlap	Figure 5-41 shows that 14 of the 33 points (24%) in the intervention phase are lower than any points in the baseline phase. There are only 2 of the baseline points which are higher than in the intervention phase (17%).

5.2.8.2 S1c – Behaviour 1 (Time Spent Refusing)

Data is presented graphically with key numerical information in Table 5-44.

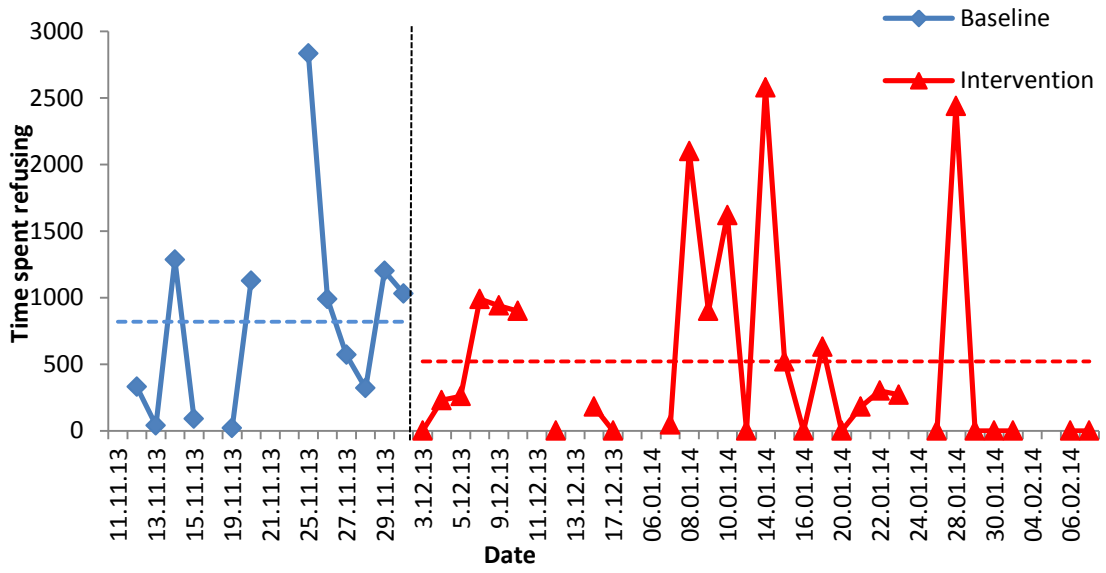


Figure 5-42: A graph showing the amount of seconds per day S1c refused, including mean lines across Phase A and B.

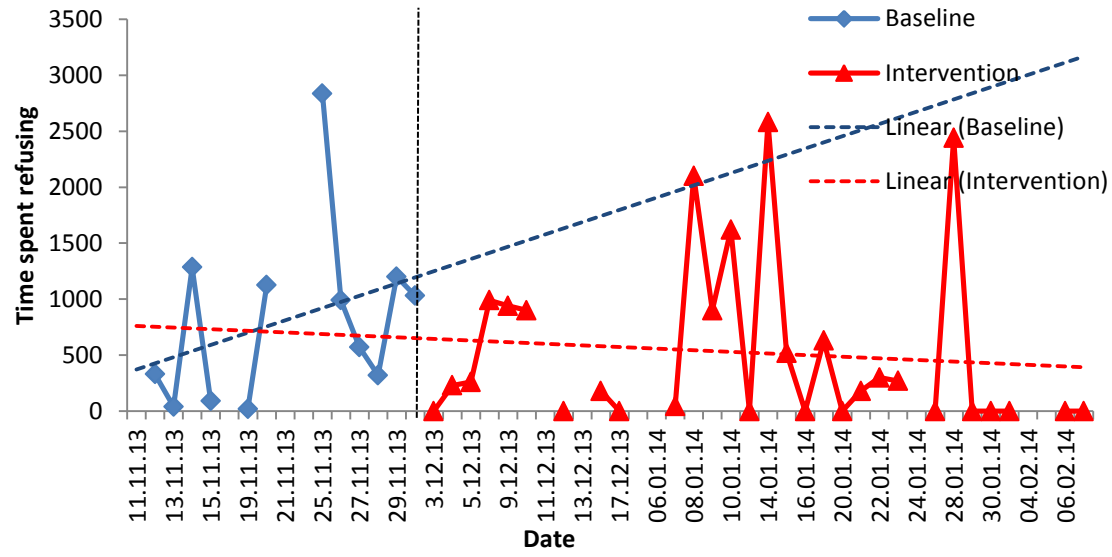


Figure 5-43: A graph showing the amount of seconds per day S1c refused, including trend lines across Phase A and B.

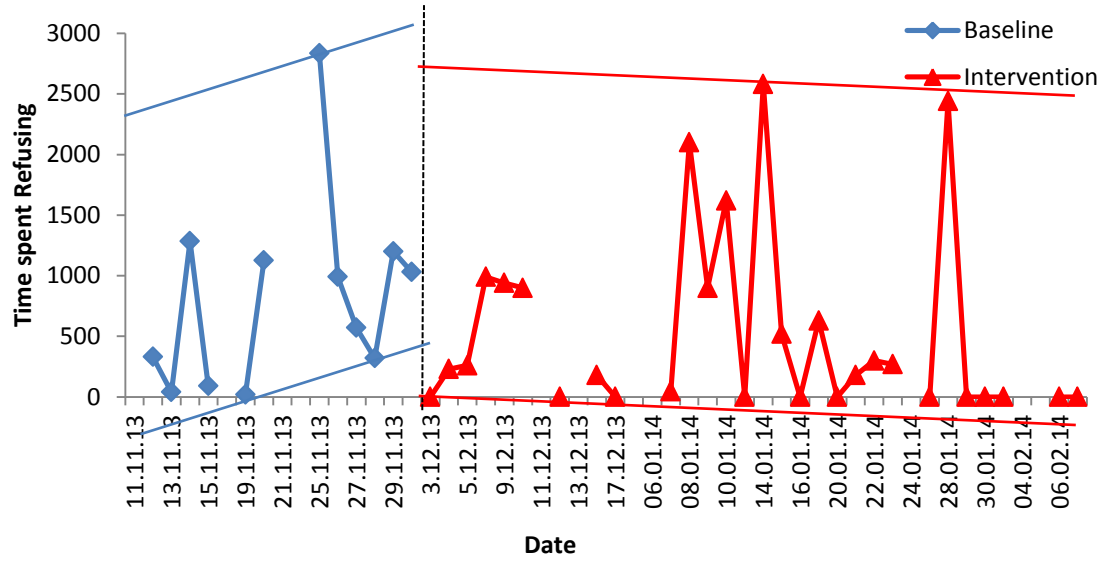


Figure 5-44: A graph showing the amount of seconds per day S1c refused, including variability lines across Phase A and B.

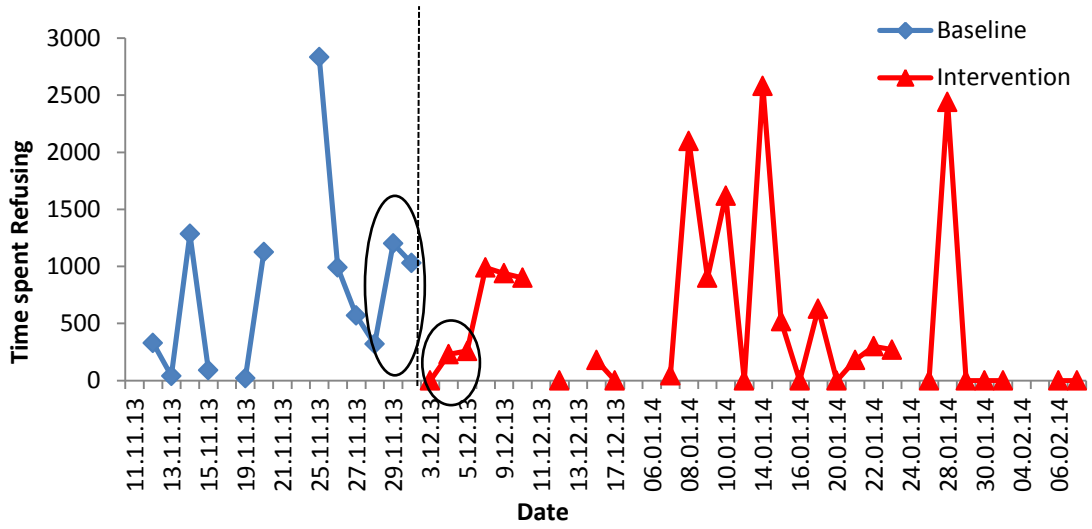


Figure 5-45: A graph showing the amount of seconds per day S1c refused, including immediacy of effect across Phase A and B.

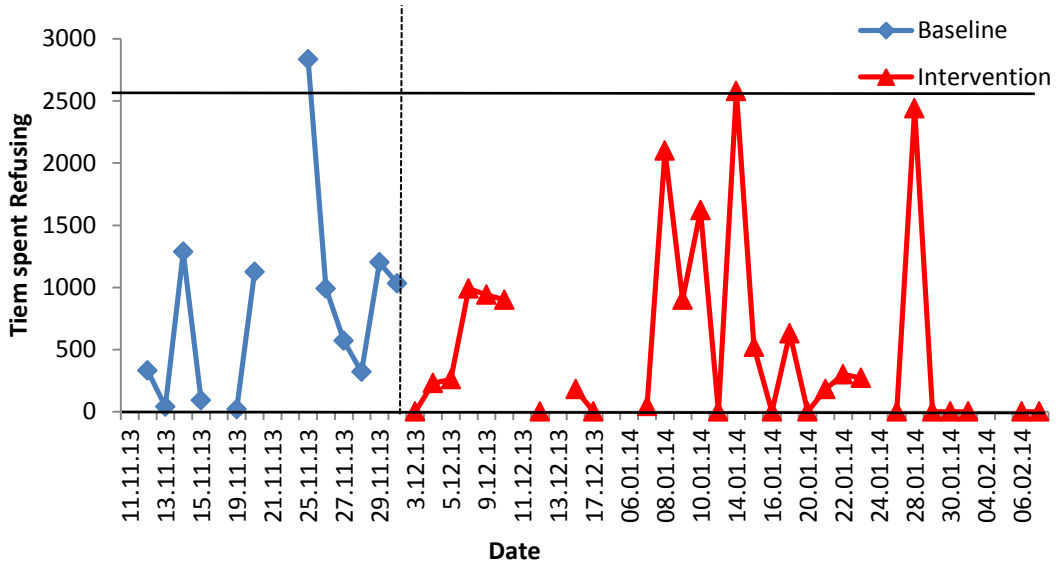


Figure 5-46: A graph showing the amount of seconds per day S1c refused, including overlap lines across Phase A and B.

Table 5-44: A table showing key numerical information for S1c (number of refusals per day).

	Baseline	**Intervention**	**Difference**
Mean (seconds)	819.58	520.17	-299.41
Standard Deviation	759.43	749.69	-9.74

Table 5-45: A table outlining brief descriptions of the outcome of the visual analysis of the amount of time spent refusing to comply each day by S1c.

Feature	**Visual Analysis**
Level	Figure 5-42 shows a decrease in the mean level (reducing by 299.41 seconds per day (Table 5-44)).
Trend	The trend lines (Figure 5-43) show that during baseline the amount of seconds per day S1c spent refusing was increased. The intervention trend line decreases over time.
Variability	The variability lines (Figure 5-44) for the intervention have slight less variation for the intervention phase than the baseline. The variability is quite large in both phases.
Immediacy of Effect	Figure 5-45 shows that the first 3 points during the intervention phase are all lower than the last 3 points in the baseline phase.
Overlap	Figure 5-46 shows that 12 of the 29 points (41%) in the intervention phase (all at 0) are lower than any points in the baseline phase. In the baseline phase 1 of the 12 (8%) points is higher than any point in the intervention phase.

5.2.9 S1c – Summary of Analysis

The number of refusals per day was clearly shown to decrease according to level, trend, variability, and percentage of overlap. The immediacy of intervention effect is unclear due to the second point being as high as the points in the baseline phase. This suggests that the intervention did decrease the number of refusals per day for S1c with a moderate to strong effect.

The number of seconds spent refusing to comply by S1c is suggested to have decreased during the intervention phase as shown by the level, trend, and overlap. The variability lines also suggest this behaviour decrease, although less clearly than the aforementioned 3 factors (see Table 5-45). The immediacy of effect does show an immediate decrease in amount of time spent refusing to comply when the intervention began, although the pattern of these first 3 points increases over time. However, considering all features together, a decrease in amount of time spent refusing to comply is observed in the data during the intervention phase with a moderate effect.

The data suggests that both the amount of times S1c refused per day as well as the amount of time spent refusing per day were decreased as a result of the intervention with a moderate to strong effect.

6 Discussion

This discussion will synthesise the results and consider these with regard to the information presented in the literature review (see 1) in order to discuss explanations for, and interpretations of, the results. Results explicitly linked to the hypotheses are discussed, as well as anecdotal evidence that explore the concepts involved in the research more broadly (in line with the pragmatic epistemology). Potential limitations of results, and overall design, are also discussed along with possible implications for future research and practice.

6.1 Summary of Findings

6.1.1 Research Question 1

1) In a sample of rural primary schools can training based on a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback mechanisms:

a) lead to use of the behaviour plan?

b) increase overall staff self-efficacy in behaviour management?

c) decrease teacher burnout?

6.1.1.1 Use of the Behaviour Plan

The results suggest that the training did not increase the amount of behaviour plans staff used 6-8 weeks later. However, of the 14 school staff who implemented the behaviour plan only 1 found the plan unhelpful. This suggests that training alone is unlikely to lead to an increase in use of behaviour plans, although it may influence which behaviour plan is implemented. 27% of staff receiving training (who completed the questionnaires at time 3) went on to implement the behaviour plan. A further 37.5% reported they were hoping to implement the behaviour plan. However, without further follow up it is unclear whether the intention to implement the behaviour plan will be actioned (see 6.2 for further discussion of this issue).

6.1.1.2 Effects of Training on Staff Self-Efficacy

The results for self-efficacy suggest that classroom management self-efficacy is not a stable construct over time (see 6.6 for other possible explanations for this instability). The measurements of change between the 3 data collection periods for the experimental group suggest some limited impact of time and training on self-efficacy. However, the complexity of interactions between training, time, and self-efficacy is made clearer when making comparisons between groups at time 1 and time 3 are made.

No significant change over time in personal or external self-efficacy was found in the experimental group scores at time 1, 2 and 3. General teaching self-efficacy was shown to increase significantly when estimated scores were included, (see 6.2, 6.5, and 6.7 for discussion of potential implications of estimated scores). This increase in general teaching self–efficacy from time 1 to time 2 was maintained 6-8 weeks after training. Overall self-efficacy also increased significantly between time 1 and 2 (and was maintained at time 3); the significant increase in general self-efficacy may have played the largest role in this increase in overall self-efficacy. However, all of the subscales suggested a trend supporting hypothesis 1b, although these trends were not significant independently, the combined effects may be reflected in the changes seen in the overall self-efficacy scores in the experimental group.

Experimental group external self-efficacy scores do not increase significantly compared to the control group between time 1 and time 3, suggesting no impact of training. Experimental group scores for personal self-efficacy did not increase significantly over time (as discussed above). However, when experimental group scores were compared to the control group, results suggest that the training may have served a protective role for personal self-efficacy, as the experimental group scores were significantly higher than those of the control group at time 3.

General teaching self-efficacy increased for both the control group and the experimental group, suggesting that this construct increases over time, regardless of training. Overall self-efficacy was shown to increase significantly when compared to the control group, which decreased. This suggests that the training protected staff from a significant decrease in overall and personal staff self-efficacy, with a potential (although not significant) slight positive impact on general and external self-efficacy (see 6.2 for discussion of potential limitations to this analysis and conclusion).

6.1.1.3 Effects of Training on Staff Burnout

As with self-efficacy, the analysis of burnout results suggest that school staff burnout measures change over time. Data for the experimental group at time 1, 2, and 3 showed no significant changes for physical and overall burnout levels. Emotional burnout increased immediately after training (the opposite effect to that predicted by the hypothesis), before then reducing back to original levels at time 3. Cognitive burnout levels decreased significantly immediately after training, these effects were not maintained at the 6-8 week follow up.

The comparison of control and experimental groups at time 1 and 3 showed an impact of training on physical, cognitive, and overall burnout in support of the hypothesis. When experimental group and control group scores were compared it appeared that training significantly decreased cognitive and overall burnout levels at follow up (the control groups burnout levels increased for these constructs). As time 2 data was not included in the comparison it is unknown if this effect was immediate. This analysis suggests training protects school staff from increased cognitive, physical, and overall burnout levels experienced by school staff over time. The results suggest that there was a potential immediate negative effect of training on emotional burnout, contrary to the direction of change predicted by the hypothesis.

*6.1.2 **Research Question 2***

2) Does use of a targeted behavioural intervention developed and implemented by a specialist behaviour service using functional analysis, solution focused principles, pupil lead target setting, and the use of regular feedback mechanisms decrease challenging behaviour in students?

6.1.2.1 The Impact of the Behaviour Plan on Student Challenging Behaviour

The results for all 3 participants included in this research suggest that use of the intervention decreased their challenging behaviours with differing levels of impact. For all 3 participants the mean level of challenging behaviours decreased notably during intervention. However, the trend lines provided a less clear result, which is likely to affect interpretation when following Kratochwill et al's (2010) analysis structure. The trend lines are likely to be impacted by the behaviour target being 0, rather than aiming for an on-going increase / decrease in behaviours over time (see 6.2 for a discussion of the impact this may have on analyses).

The data for S1a showed a decrease in "inappropriate responses for not being chosen". The decrease observed was significantly increased when the first very high data point of the intervention phase was removed from the intervention. This initial data point may have been much higher than the rest of the intervention due to an "extinction burst" (see 6.2). S1a's number of daily refusals and daily complaints were also notably reduced. S1a's "interrupting the teacher" behaviour was shown to decrease by a small amount (however, see 6.2.1.2 for discussion regarding mean levels and 0). These results suggest that the behaviour intervention had a positive impact on S1a's challenging behaviour.

The results for S1b were less clear than for the other participants. Although the number of crying incidents decreased considerably, the effect of intervention on time spent crying was not as apparent. The analysis of the time S1b spent crying was significantly increased by 2 isolated high data points in the intervention phase. According to the data sheets both of these incidents occurred during P.E. sessions (see 6.4.2). There was a clearer impact of the intervention on making "inappropriate noises" in class for S1b.

S1b's behaviours appeared to be positively impacted by the implementation of the behaviour plan, although this was less clear for "inappropriate crying" than "inappropriate noises". S1b's behaviour may have been impacted by a number of behaviour specific phenomenon such as "extinction bursts" (see 6.2). Due to sudden increases in the time S1b spent crying at specific times in the day, the intervention may not have clearly targeted functions of this behaviour (see 6.2 and 6.4.2).

The intervention appeared to have a clear positive impact reducing the challenging behaviours (refusal and time spent refusing) for participant S1c. S1c's behaviours were reduced with a moderate to strong effect. As for all participants, analysis of behaviour change for S1c was impacted by the instability of the baseline. However, for ethical reasons the intervention was implemented despite baseline instability (see 4.11 and 6.2.1.2).

The results show that the intervention seems to have benefitted all 3 students participating in the research, by reducing their challenging behaviours. Anecdotally, staff reported feeling that their understanding of the students' behaviours had increased (although not for all of S1b's behaviours). The results suggest the intervention provides a structure that can support students and staff to manage challenging behaviour. However, due to the single case design of the research, and heterogeneity of the cases and the population displaying challenging behaviours, this should be generalised with caution.

6.2 Limitations of Analysis and Conclusions.

There are a number of limitations to the data, as well as to the analysis procedures for all research questions. Although limitations of the research design and conclusion overall will be discussed in section 6.6, this section will briefly discuss limitations specific to data analysis and collection. The aim is to develop a deeper understanding of the results, before leading to a discussion of why any effects may have been / or not have been identified, and various interpretations of these results.

6.2.1.1 Research Question 1

There was a high level of dropout for research question 1. Dropout can lead to an unrepresentative sample, as confounding variables may impact continued participation in the research (see 6.4 and 6.5 for further discussion of potential confounding variables). The high dropout rate may reflect characteristics of the participants likely to complete both questionnaires, rather than staff responses to the independent variable (training). However, the dropout rate was similar for both groups, consequently suggesting if continuation in the research was dependent on specific characteristics these are similar in both groups, potentially isolating the independent variable as the only consistent difference between the staff completing the questionnaires at time 3. This dropout rate decreased the number of participants available for analysis due to the repeated measures procedures used. This was used in an attempt to minimise the impact of confounding variables (to ensure pre and post groups were equivalent). The analysis for question 1 has also been conducted on the whole group, not accounting for differences in roles or teaching experience. This is due to the independent variables intention to support school staff as a whole, leading to a group analysis. It may provide useful information to analyse the groups independently and explore differences between roles / teaching experiences. However, this was not within the scope of this research and does not directly relate to the hypothesis (see 6.4 and 6.5 for further discussion of the potential impact of these characteristics).

The data and analysis for question 1a was impacted by the low sample number, decreasing the effect size. The data for analysis of question 1a only just met most parametric assumptions and did not meet assumptions of normal distribution. Due to the

high sensitivity of the tests for normal distribution, Field (2013) recommends that the researcher considers whether or not a normal distribution would be expected for the dependent variables concerned, and view the histograms and Q-Q plots. After considering these issues, parametric tests were selected. The results must therefore be interpreted with extra caution as the tests may not be as reliable as with data sets that fully meet parametric assumptions. The data for question 1b and 1c met all parametric assumptions for most dependent variables. A small number of the dependent variables did not meet the assumptions of normal distribution. Due to the robustness of ANOVA to slight parametric assumption violations and the sample size for each group being above 30 for most dependent variables (Field, 2013), parametric tests were used as these are more reliable than non-parametric tests (Pallant, 2013). Again this means it is essential to be aware of the potential weaknesses of these analyses when the conclusions are being drawn and interpreted.

The results for question 1a are limited by the potential lack of comparability between groups, as the control group was unaware of the behaviour intervention the experimental group received training in. The question regarding intention to implement, assumes that intention to change leads to change (weaknesses of this assumption are discussed in 6.4.1).

The significance of results for general self-efficacy and overall self-efficacy altered when estimated scores were included. This is likely to have been due to the increase in participant number increasing the effect size. Although averages of the subscales were used to develop estimated scores (as the scores in the questionnaire subscales have been shown to be reliably related (Emmer & Hickman, 1991)). Conclusions from these results must be interpreted with acknowledgment of potential weaknesses of including the estimated scores, as it may be that specific questions were not answered as staff felt they were not relevant to them, rather than because they were accidently missed (see 6.5). However, adding these estimated responses did not alter the trend already shown in the actual score results, rather only increased the effect size and therefore the significance of the trends.

6.2.1.2 Research Question 2

The inter-rater reliability showed a high level (almost perfect) of agreement with regard to existence of change between baseline and intervention phases, thus increasing reliability of the analysis. Most effects found had medium to large effect sizes, according to Kratochwill et al's (2010) criteria; these are likely to be conservative due to the strict criteria and specific features of the data. The aim for the data for the behaviours is "0", meaning if the immediacy of effect is high, trend lines for the intervention phase are likely to have a shallow slope, as the ideal trend line would be horizontal at 0. This therefore results in less meaningful trend lines (decreasing effect size if the intervention trend line has less of a slope than the baseline trend line according to Kratochwill et al's 2010 criteria) than if SCEDs were completed where a continual increase in behaviours was aimed for. Consequently the mean level lines may be more meaningful for this data. Barlow et al, (2009) report that the mean level is the best indicator of behaviour change over time, especially where there is an unstable baseline. The target dependent variables being challenging behaviour also poses an issue in analysis, due to the phenomenon of the "extinction burst" (where behaviour suddenly increases when interventions are put in place), a common phenomenon when managing challenging behaviour (Cooper et al, 2007). The presence of the extinction burst can create outliers impacting on immediacy of effect, trend, and mean level lines.

Another difficulty with the analysis, is the instability of the baselines. Unstable baselines are often common in behavioural data because of the number of confounding variables in the environment (Cooper et al, 2007). Researchers introducing independent variables where there is an unstable baseline should aim for increased stability through introduction of the independent variable (Cooper et al, 2007). Although the baselines are unstable this may be representative of the behaviour itself within the environment it presents, rather than an unstable period of data collection. Awareness of this baseline instability has led to a more reserved analysis of graphs by the researcher than Kratochwill et al's (2010) criteria requires. The risk of type 2 error (Baer, 1977) is therefore further increased, leading to potential issues with unstable baselines resulting in conclusions about the effectiveness of the intervention being less than was actually the case. However, despite the potential increase of type 2 errors it has been concluded

that the intervention may be useful for managing challenging behaviour in schools, and therefore further research and application of this intervention could be helpful (see 6.7).

6.3 Additional Observations

A number of anecdotal observation opportunities and data collection procedures presented themselves during the research (see Appendix 18 for summary of information). Due to the pragmatic epistemology of the research it is important to consider all available information, not only the information gathered through pre-planned data collection procedures, to provide further insight into the explored phenomena.

The feedback from staff after training tended to be very positive. 1 staff member reported having a different view on a student's behaviour since training. Despite not implementing the plan the teacher did report having spoken to the student about the function of the behaviour and possible alternative behaviours, reporting that this led to a decrease in the challenging behaviour. This suggests that the information provided in training may have supported teaching practice, even where teachers did not fully implement the behaviour plan. A number of other staff members who did not implement the plan also told the researcher that the information provided in training was helpful. This suggests that training sessions similar to the one provided in this research may be useful for school staff, in addition to those detected through more structured data collection procedures used within this research. However, as this information was gathered anecdotally, it is very prone to selection bias.

Participating school staff were asked how many students they felt needed additional behavioural support in their classes. The average number of students staff felt would benefit from additional support at time 1 was 1.4 for both groups. At time 3 staff felt the average number of students who would benefit from additional behavioural support was 2.8 and 2.4, for the experimental and control group respectively. This suggests that despite recognising a need for behavioural support, teaching staff do not necessarily implement suggested strategies in order to support students, even after attending training on how to do so. However, the number of students with a behaviour plan in place did increase in the experimental group (from 1.7 to 2.2) and remain the same in the control group (0.7). This suggests a possible slight effect of training on behaviour plan implementation. Analysis of question 1a, however, shows these group differences

are not significant.

The teachers who were supported to deliver the intervention for research question 2 all felt the behaviour plan was partially helpful for managing challenging behaviour. However, the teacher of S1c felt that the time spent "refusing" was increasing and the number of incidents decreasing. The data, contrastingly, shows that both time and frequency were decreasing. The teacher of S1a and S1b felt that the behaviour plan was having little to no effect for both students, however, again contrastingly, the data showed a varying effect for different behaviours, including some strong positive effects. This suggests that without the clear data the teaching staff in school may not have realised the positive impact the behavioural intervention was having on the students. This rigorous data collection procedure was only completed for the purpose of the research, rather than as part of the intervention itself. This perceived underestimation of the amount of effect a behaviour plan has may impact school staff willingness to implement behaviour plans, and provide some explanation as to why behaviour plans are often not implemented even when staff recognise a need for additional support (see 6.4.2).

6.4 Impact of the Training and Intervention

6.4.1 Factors Impacting Use of the Behaviour Plan

As stated above (6.1.1.1 and 6.2.1.1) the use of the behaviour plan was recorded and explored through comparison with the number of behaviour plans in place in classrooms. The number of behaviour plans implemented was not increased after training. This suggests that training in a behaviour plan in the form it took for this research is not likely to increase the implementation of behaviour plans. However, a large portion of those implementing behaviour plans after training used the behaviour plan related to the training provided. This suggests that training providing staff with suitable behaviour plans may be helpful to encourage use of certain preferred behaviour plans. If training focuses on evidence based plans this could increase evidence based practice in classroom management of challenging behaviour. However, without knowing the types of behaviour plans implemented by control group participants and their successes it cannot be determined whether the behaviour plan from this training would have been more suitable than those which would have been implemented anyway.

The results consider intention to implement the behaviour plan. Further research (e.g. longitudinal research) would be necessary to determine whether intention to implement leads to implementation. A review of training research in schools showed a "robust" finding that individual training sessions without specific follow up are inadequate for leading to behaviour change and ensuring staff implement evidence based practices (Barton, 2013). Bubb & Earley (2013) found that training for school staff tends to be reliant on 1-2 hour sessions, given current time constraints in schools. Delivering after-school training sessions lasting for 1-2 hours are becoming increasingly common in EP practice. This highlights a need for exploring and researching factors that can increase effectiveness of training within these time constraints that leads to successful staff behaviour change (see 6.7).

Some staff reported partially implementing the behaviour plan, possibly due to the training incorporating theory, as well as a specific behaviour plan (Lang et al, 2010).

This links to anecdotal feedback from staff who reported having found the information "useful" and making them "think differently in class" after training. This increased input around theory may allow for staff to feel more ownership of their behaviour and choices to implement strategies, increasing their "locus of control", resulting in increased behaviour change (Reese et al, 1984). Due to the small sample of staff receiving direct support with intervention implementation, and having implemented the behaviour plan, it was not within the scope of this thesis to consider the impact of this additional support on self-efficacy and burnout. However, previous research has suggested that models of training that embed a coaching model of training into the school culture rather than being "owned" by the EP (e.g. the coach consult method (Balchin, Randall & Turner, 2006)) leads to increased embedded sustainable change for teachers (and students). This may highlight additional benefits of incorporating follow up sessions, negotiating specific content with staff groups, and including modelling and coaching (Balchin et al, 2006) in staff training. Macleod, Macmillan, & Norwich (2007) suggest that schools can benefit from training where select staff are trained to be trainers for the staff in their own school, which can again increase ownership of training by the school and lead to increased locus of control felt by staff, potentially leading to increased engagement (Reese et al, 1984). Little research has explored these consultative training methods (Higgins & Gulliford, 2014), possibly due to the intense time commitment required. Future research that explores these more consultative methods of training delivery for behaviour training for staff may provide further insight into how effectiveness of behaviour training can be influenced, and how these different delivery methods may impact staff.

A member of staff who did receive support to implement the behaviour plan felt the plan was highly intense and would be difficult to implement without additional support. This may indicate the behaviour plan itself is not manageable, although the teacher did clarify that the data collection procedures (taken for research purposes only) were the most unmanageable aspect (see 6.4.2). The other 3 staff members implementing the plan felt the intervention as well as data collection were manageable. Consequently, this may have reflected attributes of the particular staff member, such as low levels of self-efficacy, which often results in less commitment to the role (Lee et al, 2013) and less engagement with training and new strategies (Britten & Lai, 1998). High levels of

burnout may also result in teaching staff being too exhausted to implement the plan effectively (Farber, 1991 as cited in Jennet et al, 2003) (see 6.4.3).

6.4.2 Impact of Intervention

The intervention was implemented fully by 13 members of staff who received training. A further 3 participants reported having partially implemented the behaviour plan. 13 of the 16 using the behaviour plan felt the intervention had a positive effect on the students. All of the staff taking part in research question 2 felt the intervention helped increase their understanding of the student's behaviours. Increased staff understanding of behaviours is likely to result in a more positive reaction to the students (Mossman et al, 2002) which in turn can create a more empathetic and nurturing environment, potentially creating a more positive and success based learning environment for the student (Mitchell & Hastings, 2001).

The analysis of research question 2 suggests that use of the intervention for 6-7 weeks was successful in reducing challenging behaviours for all 3 participants (although to differing degrees for varying behaviours). The majority of staff additionally reported that they felt implementation of the behaviour plan was manageable (although a number of staff reported feeling it seemed unmanageable). Those receiving support to implement the intervention were also required to take daily data during the research period. The teaching staff felt that data collection was not maintainable and made the intervention much more difficult to implement in the classroom. The data collection procedure was for research purposes only, not the intervention itself. However, staff implementing the intervention felt the intervention was not / was having a very small impact on challenging behaviours. This underestimation of behaviour plan impact could risk teachers stopping potentially helpful interventions. Although behaviour rates decreased, if teachers did not feel the behaviour had decreased the intervention may not be determined successful (when defining behaviour as a social construction (Emerson, 2001)), unless their perceptions of the student exhibiting challenging behaviour altered. Implementation of the behaviour plan may increase burnout levels due to the increased additional effort and time commitments involved (see 6.4.3). Alternatively, challenging behaviour may be so demanding on teaching staff (DfE, 2012a) that changes must be

dramatic before staff perceptions alter, possibly decreasing the perceptions of benefit of new interventions compared to time costs for teachers. However, the results suggest that a targeted behavioural intervention using functional analysis, solution focused principles, pupil led target setting, and the use of regular reinforcement can support staff to decrease student challenging behaviours. The use of this plan also reportedly increases understanding of behaviours, which may in time increase staff positive affect (Mossman et al, 2002), leading them to appreciate the behaviour change over time. A decrease in challenging behaviours is likely to impact the target student's educational experiences, the other students in the class (DfE, 2012a) and the teacher's own health (Hastings & Bham, 2003) (possibly including self-efficacy and burnout see 6.4.3).

Students spoke positively about the intervention both to the researcher and to the teachers they were working with. This may be due to the solution focused and student led aspects of the behaviour plan encouraging them to create a more positive version of themselves (Rae & Daly, 2008). Previous research has been sparse and unclear as to the effects of solution focused work (Kim, 2008), target setting, and behaviour monitoring in young children. However, this research suggests that giving students (even of a young age) some "locus of control" (Reese et al, 1984) may be helpful when combined with regular rewards and feedback. This research further suggests that an intervention combining these solution focused and pupil centred principles with the principles of FBA and DRO may be helpful for managing the challenging behaviours of primary school students.

McIntosh et al (2008) have shown that FBA is successful in altering behaviours. FBA has been shown to be more effective for maintenance and generalisation when combined with strategies that consider additional systems (Evans et al, 2003). It may be that due to the multi-dimensional element to this behaviour plan, effects will be maintained over time. However, due to the timescale of this research, maintenance effects have not been recorded.

The behaviour plan was less successful for participant S1b for the amount of "time spent crying". This may be due to approaches based on functional assessments and

reinforcers ignoring emotional needs (Winnet & Winkler, 1972). However, through the multi-dimensional approach of the behaviour plan, it should be possible to adapt the behaviour plan to ensure it is personalised to the individual, focusing on the aspects of behaviour plan that are more suited to the particular behaviour or environment of the individual each week. The data highlighted patterns in crying time (increased during P.E.) which may guide understanding of behavioural functions. S1b did not select crying as a target behaviour during the research period. This may be because S1b did not feel "crying" was a challenging behaviour, or due to S1b not feeling adequately emotionally supported to decrease this behaviour. As crying is likely to be attached to emotional outcomes for a student, it is possible that S1b was not able to manage the emotional repercussions of acknowledging and discussing this behaviour (S1b was very reluctant to discuss this behaviour during the sessions observed by the researcher). The solution focused aspect of the behaviour plan may encourage S1b to develop a positive self-view (Johnson & White, 1971) during the intervention, potentially leading to a view of himself where he can manage his crying. This research suggests that the theories included in this intervention appear to be an effective way of guiding younger students through the solution focused process (Elliot & Faupel, 1997). The small but visible decrease in time spent crying, despite no targeted work on this behaviour, does suggest that S1b's crying was impacted by the changes he was implementing. The rate of change may be decreased as it relies on secondary effects rather than primary.

Data from research question 1 and research question 2 suggest the intervention was helpful in decreasing the challenging behaviour of the primary school students in the SCED and whose teachers responded and implemented the behaviour plan independently. Only 1 teacher independently delivering the intervention felt it was not helpful for the student. Although data had not been taken (to the researcher's knowledge) on the behaviours of the students whose teachers implemented the behaviour plan, but were not involved in research question 2, these staff reported feeling the intervention was helpful. If adhering to the socially constructed definition of challenging behaviour (Emerson, 2001) the behaviour plan could therefore be concluded as successful. Teachers' belief that they are able to successfully decrease challenging behaviours of students may decrease their burnout (Eyged & Short, 2006) and increase their self-efficacy as they experience successful implementation of

strategies (Dellinger et al, 2008). This is likely to improve student outcomes, and potentially improve staff health (Hastings and Bham, 2003) and confidence, increasing their commitment to teaching (Ware & Kitsantis, 2007).

6.4.3 Impact of Training on Self-Efficacy and Burnout

The teacher outcomes of training on school staff self-efficacy and burnout were varied according to individual subscales within the overall constructs. This could potentially be due to a number of confounding variables impacting self-efficacy and burnout measures, such as leadership and collective staff efficacy levels within schools (Kurt et al, 2012). Although allocation of schools was randomised to minimise effects of school culture, school culture may still explain why different staff valued the training differently. In some schools teachers commented that the training was "no different" (questionnaire response) to what was already in place, however when speaking with senior staff from these schools they tended to report that the training covered new theories and interventions. This may reflect an unawareness of staff practice by senior leadership (possibly resulting in school culture impacting on self-efficacy and burnout measures), or it could reflect the self-efficacy and burnout levels of these staff members themselves. Klingner et al (2003) found that staff are less able to engage in training and less willing to implement new strategies if they are already burned out or have low self-efficacy. Consequently, it may be that staff members who felt they did not learn "new" information at the training were unable to engage with the training. This highlights a potential need for staff self-efficacy to be increased and burnout to be decreased before conducting training in schools, so that they are better able to benefit from the information shared (see 6.9). These confounding variables may be reflected in the finding that self-efficacy and burnout measures did not appear to be stable over time. Some of the positive effects of training found within this research concerning staff burnout and self-efficacy were only observable when compared to the control group. The training appeared to serve a protective role for staff burnout and self-efficacy (reducing negative changes over time) rather than creating significant within group changes at pre and post measures.

The experimental group showed no change in personal or external self-efficacy over

time, but did show a significant increase over time in general and overall self-efficacy scores. However, when compared to the control group it appears general self-efficacy increases over time in both groups, suggesting that experience may impact general self-efficacy rather than training. As the training was conducted near the beginning of the academic year, the general teaching self-efficacy of the staff may have increased as familiarity with their students increased, and they had an increased number of direct mastery experiences with their class (Bandura, 1977). The results suggested a protective function of training in managing challenging behaviour on personal self-efficacy. This may be due to increasing staff knowledge as a result of training, and staff consequently feeling they are better equipped to manage challenging behaviour. Newman–Carlson & Horne (2004) showed a similar effect on school staff after training in bullying prevention, where increased knowledge led to increased belief in their ability to manage these specific situations only. This suggests that interventions aiming to support self-efficacy may need to be domain specific (Bandura, 1977). A number of studies presented in the systematic review (see 1.4) showed an increase in the self-efficacy of teaching staff but did not report individual subscale scores (e.g. Palmer, 2010; Stoiber & Gettinger, 2011). The increase in overall self-efficacy in classroom management after training, compared to controls in this research may reflect the apparent protective mechanism of training found on personal self-efficacy alone, although all domains suggested a trend in support of the hypothesis, which when combined may have impacted results sufficiently to create a positive impact on overall staff self-efficacy. This increase in personal belief in one's ability to cope is likely to impact utilisation of skills (Bandura, 1977). Training that increases knowledge and makes explicit links to theory (Lang et al, 2010) may give the "locus of control" to staff (Reese et al, 1984), increasing staff self-image and belief that they can alter behaviour through application of strategies.

Effects of training on teaching staff burnout levels has received less attention in the literature than self-efficacy, despite research suggesting the two constructs are correlated (Eyged & Short, 2006), as well as highlighting connections between burnout and staff commitment (Ware & Kitsantis, 2007). The current research considered impact of training in challenging behaviour on staff burnout. Again, the scores in the control group suggest that these factors are changeable over time, and confounding variables

similar to those discussed for self-efficacy may play a role in these changes.

Emotional exhaustion was shown to increase immediately after training in the experimental group. This may be due to staff being made more aware of strategies they could be implementing and feeling additional pressure after training. The training may also make staff more aware of challenging behaviour in the classroom, possibly increasing negative reactions and therefore increasing burnout levels (Mitchell & Hastings, 2001). As training may encourage staff ownership of the problem, a risk of increased burnout arises as training may lead staff to feel more pressure after training to manage the behaviour, rather than attribute behaviour to external forces (Brouwers & Tomic, 2000). However, emotional exhaustion was shown to decrease again by follow up and no difference was found between control and experimental group scores at time 3, suggesting any negative effect of training on emotional exhaustion was short term. This is contrary to Hall et al's (1977) research showing training in classroom management decreased emotional exhaustion by increasing the sense of personal accomplishment. Hall et al's (1977) study may have led to these results as the training was more intensive and long-term, possibly supporting staff to experience success in strategies and application of knowledge rather than leaving them to apply strategies independently. This may suggest that in order to support emotional exhaustion in teaching staff, long-term training is necessary (Barton, 2013).

Cognitive burnout decreased immediately after training which may be due to the staff being provided with additional knowledge. Impact of training on cognitive burnout levels were not maintained 6-8 weeks after training. However, cognitive burnout levels in the control group increased significantly over time (time 1 to time 3) compared to those who had received training. This suggests training appeared to play a role in protecting staff from cognitive burnout, the same effect was found for physical and overall burnout levels over time.

Previous research has suggested that, in order to alter staff behaviour and attitudes, one-off training sessions are not sufficient (Barton, 2013). Staff provided different feedback about whether or not they felt the intervention was applicable and manageable in

classroom settings. This is possibly due to differing individual initial levels of staff burnout and self-efficacy, potentially resulting in staff feeling differently about their capacity to implement strategies (Klingner et al, 2003), thus leading to a decreased impact of training for the most vulnerable staff (although pre-tests did not highlight outliers and showed initial group scores were not significantly different from one another). This theory may suggest that more change for vulnerable staff would be created if self-efficacy or burnout was a direct focus (Palmer, 2010), followed by domain specific training aimed at altering behaviour. Britten & Lai (1998) showed that as self-efficacy increases, time engaged with training also increases, leading to increased positive impact on staff knowledge and self-efficacy, supporting this theory.

With the data gathered it was impossible to determine whether staff who received on-going support to implement the training had increased self-efficacy and decreased burnout compared to those receiving training without further support. However, this on-going focused support is reported to be more likely to elicit school staff behaviour change (Barton, 2013), which may, in turn increase self-efficacy and decrease burnout as staff are supported to develop their practice and experience success (Bandura, 1977). Overall the results suggest that individual training sessions in challenging behaviour can have a positive impact on staff self-efficacy and burnout levels. This extends the findings from the systematic review (see 1.4), which demonstrate that prolonged training can have a positive effect, to suggest that even isolated training sessions can have a supportive role for teaching staff in regards to these outcomes.

6.5 School Culture

A number of issues were highlighted during the course of the research that were not directly explored through the hypotheses (see Appendix 18). A number of these outcomes seemed to be related to school culture in regards to views of research in schools, as well as the stark differences between the way staff in various positions perceive their roles regarding challenging behaviour.

School SENCOs / head teachers or lead behaviour teachers were given the pre-training questionnaires prior to the researcher meeting the staff. On a number of occasions the researcher was contacted to "check" whether TAs should also be asked to complete the questionnaires. A number of senior staff (including a number of TAs) reported being unsure whether the questionnaires were relevant to TAs. This was usually due to the question about how many students in the staff member's class they felt would benefit from additional behavioural support, and the self-efficacy questions about confidence with managing various difficulties in the classroom. This suggests that although TAs are often involved in the day to day running of the classroom they are not necessarily perceived by themselves or others to be able to comment on behaviour management. This may not only impact their willingness to partake in research as they do not feel their contributions are valid (possibly accounting for the large dropout rate of TAs), but also on factors about their views of themselves (potentially related to self-efficacy and burnout). This effect may also have impacted the high level of incomplete self-efficacy forms at time 1 (however, when the printing errors in the questionnaires were corrected for time 2 and 3 this effect was no longer present).

Staff dropout was an issue for most schools in both the experimental and the control group. In a few schools teachers gave feedback (either written on the forms or through the contact staff member) that they did not have time to complete questionnaires 2 or 3 times. Despite staff being informed that the research was aiming to inform future training and use of the behaviour plan within their school and local area, they were often not committed to informing the research. The researcher received a report from 1 staff member that as a teacher they were too busy to fill in forms "not relevant" to their jobs. This may reflect a lack of training about evidence based practice for school staff, a lack

of interest, or possibly the time constraints and strict outcomes expected from teaching staff. This lack of time and outcome related pressure may not only reduce willingness to dedicate time to research, but is also likely to impact on self-efficacy and burnout levels, possibly decreasing staff effectiveness (Farber, 1991, as cited in Jennet et al, 2003).

This feeling of lack of time to complete the measures may also be reflected in the feedback given by some staff that the intervention was too time-consuming to implement. Pressure and time restraints felt by staff may also influence their feelings that outcomes of the behaviour plan would not outweigh the costs of implementing the plan (see 6.4.1). This may however, also be impacted by issues such as high staff burnout, and / or low staff self-efficacy resulting in staff feeling they do not have the capacity to implement additional strategies (Ware & Kitsantis, 2007) despite potential benefits.

6.6 Research Limitations

The current research has a number of limitations, these are discussed in this section so that conclusions about the research outcomes can be fully informed. The limitations of this particular research can also be useful in providing information for avenues for further research. As a number of limitations specific to the data and analysis have been previously discussed (see 6.2), this discussion will focus on design and implementation limitations.

The real world nature of this research immediately presents difficulties, as conditions cannot be as controlled as in strict laboratory experiments (Robson, 2011). The current research was susceptible to a number of weaknesses common in real world research such as high rates of participant dropout, missing data points (due to staff being unavailable / unforeseen changes in the classroom), and maintenance of staff commitments (Maruyama & Deno, 1992). Difficulties regarding fidelity checks of data taking procedures were created as the presence of the researcher appeared to impact student behaviour (with 0 presentations of behaviour when the researcher was present). The decision was consequently made to stop these fidelity checks after the initial 2 checks for each participant. Although stopping fidelity checks minimises observer effects on the students, it did increase risk of inaccuracies of data collection. The environment of classrooms and changing staff / student number would have made it extremely difficult and time consuming (and potentially impossible) to ensure that any data collection fidelity checks accurately reflected a whole week of data collection.

Fidelity checks were put in place every second week for the intervention itself (Appendix 15) which increased reliability of the independent variable. However, there were no fidelity checks completed for the staff training (as this was delivered by the researcher), possibly resulting in certain staff receiving different information at training as the researcher became more familiar with the materials. However, the nature of training given to schools by EPs is often flexible according to each school's needs, increasing the ecological validity of the training provided as part of this research.

The design of research question 1 did not account for the difference in schools where some individual staff were being supported to implement the behaviour plan. This additional support may have served as a confounding variable for some of the data. However, as only 4 members of staff received direct support, this is unlikely to have had any significant impact on whole group outcomes. Their data was not highlighted as an outlier within the group results, suggesting the additional support given did not significantly impact the group dependent variables.

Consideration of impact of training on self-efficacy and burnout over time was limited as time 2 data was only available for the control group. A clearer picture of the impact (especially as it appears these factors were not constant over time even within the control group) would have been gained had data been available for both groups at all time points. However, due to pragmatic difficulties and staff commitment to completing questionnaires (see 6.5) this did not seem possible.

The data gathered for research question 1 was in the form of questionnaires. The results of the questionnaires showed self-efficacy and burnout as unstable over time (see 6.1.1.2 and 6.1.1.3). This may be a result of the questionnaires not measuring the construct which it claims, but being susceptible to confounding variables such as participant mood, or environmental factors. However, as the constructs are shown to be related to and affected by environmental and work stressors rather than only within person constructs (Maslach, 2003), and dependent on experiences (Bandura, 1977), one would expect the constructs to change over time as teachers experiences and work-loads alter. Consequently the changing self-efficacy and burnout levels found in the control group may be a reflection of changes teachers tend to experience at different stages in the academic year.

Questionnaires are inherently prone to flaws such as response bias, desirability effects, and subjectivity of responses. Questionnaire studies are especially vulnerable to desirability effects where the researcher is known to participants, or where the researcher delivers the intervention (Kaspereen, 2012). Consequently results may have been impacted by this, although in order to attempt to minimise impact, participants

were regularly reassured of the anonymity of analysis procedure. Questionnaires are also commonly at risk of selective sampling resulting in participants with certain traits / experiences replying, skewing the results to represent this particular sample rather than the whole sample targeted. Within this research this may be the case as there was a high level of participant dropout, which was particularly high for TAs. This suggests that the results should be interpreted with additional caution when generalising effects to TAs. It was not within the scope of this thesis to consider data separately according to role. However, this may have decreased the issue of selectivity bias as some confounding variables could have been accounted for within the analysis. Using repeated measures analysis procedures aimed to minimise these effects by ensuring participants were compared to their own results across time points.

The use of SCED raises a number of limitations especially the use of an AB design, which is not as reliable as multiple baseline or withdrawal designs (Kratochwill et al, 2010). Due to the nature of the intervention and behaviours a withdrawal design would not have been possible or ethical, and a multiple baseline design would not have been possible given the timescale. Use of a multiple baseline design would have resulted in a number of students not receiving the intervention for a prolonged baseline, which due to the behaviours being targeted may have been ethically inappropriate. The nature of the behaviours being targeted (harmful to the education of the participants and potentially other students) resulted in the decision that the intervention should be implemented despite stable baselines not yet being established, for ethical reasons. This results in increased difficulties when drawing conclusions from the data (see 6.2.1.2) due to fundamental SCED principles requiring the evaluation of baseline trends to understand the natural patterns of the targeted behaviour (Barlow et al, 2009). Due to the individual focus of SCEDs they are limited in their ability to be generalised to the population. They can however, provide insight into "what works with whom" (Pawson & Tilley, 1997). Due to the limited sample number and heterogeneity of the sample, further research would be needed with different ages, behaviours, school settings, and with larger samples before generalisations should be made.

6.7 Impact on Education Practice

The research not only highlights various areas for future research, aiming to influence and support teaching staff in order to manage challenging behaviour (see 6.9), but also leads to ways in which educational professionals may be able to begin doing so.

The intervention on which the training was based appeared to have a positive effect on the management of challenging behaviours in class (see 6.1.2.1). Consequently, it is likely to be beneficial to both classroom staff and services responsible for training and delivering such interventions to guide ways in which staff can be supported to implement plans based on the psychological theories within this behaviour plan. Training teachers to implement behaviour plans and apply these theories independently may reduce referrals for more specialist services (e.g. in the county where the research was based, prior to this research the BEMS was required to implement the behaviour plan, limiting the number of pupils accessing it), therefore allowing them to work in a more consultative manner to support more staff and students. Implementation of plans may improve teacher knowledge of their students and student behaviours, increasing their positive affect towards students, and therefore student success (Mossmann et al, 2002).

Research question 1 focuses on the outcomes of 1-2 hour staff training sessions. The findings suggest that some staff find this useful and supportive, and it can impact the type of behavioural interventions that are implemented. However, despite training, some staff still felt they had a number of students in their class who would benefit from additional support without behaviour plans. Consequently, it seems 1 training session may not be enough to impact on staff behaviour in a number of cases. Although elements of self-efficacy and burnout improved after training, some of these changes were not maintained. Nevertheless, a number of staff did report they felt the training was useful, despite not having implemented the behaviour plan. It is therefore possible to conclude that individual training sessions do provide some support to staff, benefitting their self-efficacy and burnout, and consequently their health, commitment, and ability to fulfil their teaching roles (Ware & Kitsantis, 2007). However, in order to influence specific actual behaviour change (rather than just motivation to change) this

research highlights a need for educational professionals to provide staff with training staff that includes additional long-term support or more regular training sessions (Barton et al, 2013).

Additional findings referred to in this research as issues relating to "school culture" (see 6.5 and Appendix 18) suggest there may be a need to clarify TA roles, while supporting them to feel they have valid contributions to make in regards to student learning and behaviour. These additional observations also suggest that staff may need to be supported in understanding the relevance of educational research to their work, so that they feel contributing their time to research is a valid and helpful part of their role.

6.8 Unique Contribution

The research concerning training sessions in literature does not tend to consider the impact of individual training sessions. The literature claims that for valuable change to take place sessions must include long-term and follow up plans (Barton et al, 2013). However, the current research demonstrates that single training sessions covering challenging behaviour theory and intervention can lead to some positive behaviour and attitude change in school staff. This is especially important given the current economic climate which exerts increasing pressures by limiting time for training in schools (Bubb & Earley, 2013). The current climate change concerning education in the UK highlights a need for successful multi-agency working. This research aimed to consider ways in which professionals (EPs and specialist teachers) can work together to effectively disseminate training.

The behaviour intervention was used with primary school aged children, to the researchers knowledge there has been no previous research determining whether interventions based on solution focused and student led principles can be effective with younger students. This research suggests there is a potential for solution focused and student led behaviour plans to be helpful for primary school aged children, which is especially important given the new governmental focus on student voice (Children and Families Bill, 2013.

This research explicitly considers impact of training on challenging behaviour (a need that is felt by teaching staff, (NFER, 2012)) on self-efficacy and burnout. To the author's knowledge only 1 other study has considered this (Stoiber & Gettinger, 2011) which involved long-term training. Therefore this research may be more ecologically applicable to the current school climate (Bubb & Earley, 2013) as less time is required from specialist services to deliver training than previous research (Stoiber & Gettinger, 2011). Although the research cannot definitely state that individual training sessions can change behaviour and attitudes, it does suggest that individual training sessions can be helpful for school staff (and therefore ultimately students).

6.9 Further Research

As discussed previously there are a number of limitations within this research which highlight avenues for further research. Further exploration of the issues within this research may be useful in understanding the factors influencing challenging behaviour management, staff burnout, and staff self-efficacy, as well as other related constructs. Generalisability is an issue for both research questions, and in order to generalise findings about the impact of the intervention it may be helpful to involve further students of different ages, with different challenging behaviours, and in different settings. Further SCED research may highlight patterns around "what works for who" (Pawson & Tilley, 1997). This may then be used to guide group methodologies to further determine causal factors and impact of the intervention. Comparisons between different interventions could also begin to determine whether training in multi-dimensional behavioural interventions and psychological theories (as used within this research), or whether applying psychological theories and strategies from individual psychological theories is more useful for staff and students.

Due to the small sample of staff receiving direct support with intervention implementation, and having implemented the behaviour plan, it was not within the scope of this thesis to consider the impact of additional support on self-efficacy and burnout. Future research exploring whether training alone is sufficient, or whether additional support is required after training to result in staff behaviour change may therefore be useful. Exploration into whether training of the behaviour plan alone, theory alone, or both theory and plan is most supportive, may help to determine any individual contributions of these aspects of training that contributed to findings within this research. The type of trainer giving the training may also impact the outcomes and this may be different according to the roles of the trainees. As discussed above (see 6.5) the dependent variables may have been impacted by staff role, and further exploration may determine what type of training packages are most effective for different staffs' well-being and behaviour change. Due to the limited representation of different roles within the final sample, the "role" was not considered, despite (as discussed in relation to school culture (see 6.5)) the possibility that this may impact a number of outcomes, such as engagement in research, self-efficacy, and burnout measures.

As found in the analysis it appears that self-efficacy and burnout change over time. Consequently it may be helpful to explore factors that are responsible for these changes, this exploration may further inform how these changes may impact ability to manipulate these changes / outcomes (e.g. through informing type or timing of training according to other extraneous variables leading to self-efficacy and burnout scores, or supporting staff by directly manipulating these other variables) to best support school staff. It may be that self-efficacy and burnout levels before training impact the way in which training is received by staff, therefore impacting training outcomes. Exploratory research into the factors influencing these outcomes may be able to better inform future research as to how to support staff in managing challenging behaviour in the classroom, and improve their own self-efficacy and burnout levels.

7 Conclusion

The research has shown that implementation over 6 weeks of an intervention using functional analysis, solution focused principles, pupil led target setting, and the use of regular feedback and reinforcement mechanisms can be supportive for reducing challenging behaviour in selected primary school pupils. Although the research is limited in its generalisability due to the design (SCED), and the focus on externalised behaviours, the research suggests interventions such as the one implemented in this research have potential to provide classroom staff with effective strategies and skills for reducing challenging behaviour.

Further to this, the research also suggests that individual training sessions including theory and explicit links to practical application, such as the one used in this research can have some positive impact on school staff self-efficacy and burnout levels. Although longer-term training plans may achieve greater behavioural change, or effects that are maintained over longer time periods (Barton et al, 2013), this research suggests that individual training sessions can be helpful for staff (which is especially important given the current time pressured environment in schools).

This research has highlighted a number of areas for further research that could support generalisation of findings, and potentially increase understanding of the mechanisms behind staff and student changes. However, the research suggests that multi agency working between education services (TEP and a BEMS teacher) can lead to development of useful whole school training that can have some positive impact on staff self-efficacy, staff burnout, and pupil behaviour. In this research this whole school training was used to disseminate information to a wider audience than when services work on a case by case basis, aiming to enable staff to manage behaviours in the classroom more independently. The information from the training was shown to support staff to decrease student challenging behaviour, and to have a positive impact on staff self-efficacy and burnout outcomes, these constructs have previously been related to health and job commitment.

8 References

Artman-Meeker, K., & Hemmeter, M.L. (2013). Effects on Training and Feedback on Teacher's use of Classroom Preventative Practices. *Topics in early Childhood Special Education*. **33**(2), 112-123.

Baer, D.M. (1977). "Perhaps it would be better not to know everything". *Journal of Applied Behavior Analysis*. **10**(1), 167-172.

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioural Change. *Psychological Review*. **84**(2), 191-215.

Barlow, D.H., & Hersen, M. (1984). *Single Case Experimental Designs. Strategies for studying behaviour change (Second Edition).* Pergamon Press Inc: Oxford.

Barlow, D.H., Nock, M.K., & Hersen, M. (2009). *Single Case Experimental Designs. Strategies for studying behaviour change* (Third Edition). Allyn & Bacon: Boston.

Barton, E.E., Chen, C-I., Pribble, L., Pomes, M., & Kim, Y-A. (2013). Coaching Preservice Teachers to Teach Play Skills to Children with Disabilities. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for exceptional Children*. **26**(4), 330-349.

Beadle-Brown, J., Hutchinson, A., & Whelton, B. (2012). Person-Centred Active Support – Increasing Choice, Promoting Independence and Reducing Challenging Behavior. *Journal of Applied Research in Intellectual Disabilities*. **25**, 291-307.

Balchin, N., Randall, L., & Turner, S. (2006). The Coach Consult Method: A model for sustainable change in schools. *Educational Psychology in Practice*. **22**(3), 237-254.

Brinson, D., & Steiner, L. (2007). *Building Collective Efficacy. How Leaders Inspire Teachers to Achieve. Issue Brief: The Centre for Comprehensive School Reform and Improvement*. Centre for CSRI Issue Brief.

Britten, P., & Lai, M.K. (1998). Structural Analysis of the Relationships among Elementary Teachers' Training, Self-Efficacy, and Time Spent Teaching Nutrition. *Journal of Nutrition Education*. **30**, 218 - 224.

Brouwers, A., & Tomic, W. (1999). Teacher Burnout, Perceived Self-Efficacy in Classroom Management, and Student Disruptive Behaviour in Secondary Education. *Curriculum and Teaching*. **14**(2), 7-26.

Brouwers, A., & Tomic, W. (2001). The Factorial Validity of Scores on the Teacher Interpersonal Self-Efficacy Scale. *Educational and Psychological Measurement.* **61**(3), 433-445.

Bryman, A. (2006). Integrating quantitative and qualitative research: how is it done? *Qualitative Research.* **6**(1), 97-113.

Bubb, S., & Earley, P. (2013). The Use of Training days: finding time for teachers' professional development. *Educational Research*. **50**(3), 236-248.

Buie, E. (2009). *Pupil behaviour has 'improved'. TES. 27th November 2009*. Accessed via http://www.tes.co.uk/article.aspx?storycode=6028652 on 18.03.2013.

Cameron, R.J. (1998). School Discipline in the United Kingdom: Promoting classroom behaviour which encourages effective teaching and learning. *School Psychology Review*. **27**(1), 33-44.

Camparo, J. (2013). A geometrical approach to the ordinal data of Likert scaling and attitude measurements: The density matrix in psychology. *Journal of Mathematical Psychology*. **57**(1-2), 29-42.

Carifio, J., & Perla, R. J. (2007) Ten Common Misunderstanding, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. *Journal of Social Science*. **3**(3), 106-116.

Chan, D.W. (2005). Counselling values and their relationships with self-efficacy among Chinese secondary school teachers in Hong Kong. *Counselling Psychology Quarterly*. **18**(3), 183-192.

Clerici, R. (2008). Knowledge and attitudes of future schoolteachers in the scientific-mathematical sphere: some evidences for Italy. *Educational Studies*. **34**(4), 277-287.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education*. 7th ed. Routledge: Oxon.

Colman, A. (2006). *Oxford Dictionary of Psychology*. Oxford University Press: UK

Corcoron, J. (2006). A Comparison Group Study of Solution-Focused Therapy versus "Treatment-as-Usual" for Behaviour Problems in Children. *Journal of Social Service Research*. **33**(1), 69-81.

Cooper, J.O., Heron, T.E., & Heward, W.L. (2007). *Applied Behavioural Analysis*. 2nd Ed. Pearson Education Inc: New Jersey

Crone, D.A., Hawken, L.S., & Bergstrom, M.K. (2007). A Demonstration of Training, Implementing and Using Functional Behavioural Assessment in 10 Elementary and Middle School Settings. *Journal of Positive Behavior Interventions*. **9**(15), 15-31.

Daki, J., & Savage, R.S. (2010) Solution-Focused Brief Therapy: Impacts on Academic and Emotional Difficulties. *Journal of Educational Research*. **103**(5), 309-326.

Dawson, P., & Guare, R. (2010). *Executive Skills in Children and Adolescents: A Practical Guide to Assessment and Intervention*. The Guilford Press: London.

de Courcy-Bower, L. (2011).The extent to which collaborative teams of educators link the results of functional assessment to function-based interventions. *Dissertation Abstracts International Section A: Humanities and Social Sciences*. **71**(8-A), 2746.

Dellinger, A.B., Bobbett, J.J., Olivier, D.F., & Ellett, C.D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS – Self. *Teaching and Teacher Education*. **24**(8), 751-766.

Department for Education and Employment. (1994). *Code of Practice for the Identification and Assessment of Special Educational Needs.* London: DfEE.

Department for Education (DfE). (1994). *The Education of Children with Emotional and Behavioural Difficulties. Circular **9**(94),* London: DfE.

Department for Education and Skills (DfES). (2001). *Special Educational Needs: Code of Practice.* Nottinghamshire: DfES Publications

Department for Education. (2010). Pupil behaviour in schools in England. Education Standards Analysis and Research Division. *Research Report DFE*-RR218.

Department for Education. (2012). Pupil behaviour in schools in England. Education Standards Analysis and Research Division. *Research Report DFE* RR-218

Department for Education. (2012b). *Ensuring Good Behaviour in Schools*. *Education Standards Analysis and Research Division*. Crown Copyright, 2012.

Department for Education. (2013). *National Statistics: Permanent and fixed period exclusions from schools in England: 2011 to 2012 academic year*. Department for Education: London.

de Shazer, S., & Berg, I. K. (1997). What works? Remarks on research aspects of solution focused brief therapy. *Journal of Family Therapy*. **19**, 121−124.

Dunlap, G., Strain, P.S., Fox, L., Carta, J.J., Conroy, M., Smith, B.J., Kern, L., Hemmeter, M.L., Timm, M.A., McCart, A., Sailor, W., Markey., U., Markey, D.J., Lardieri, S., & Sowell, C. (2006). Prevention and Intervention with Young Children's Challenging Behavior: Perspectives Regarding Current Knowledge. *Behavioral Disorders*. **32**(1), 29–45.

Elliott, H., & Faupel, A. (1997). An interpersonal problem-solving approach. *Educational Psychology in Practice: theory, research and practice in educational psychology*.**13**(1), 21-28.

Emerson, E. (2001). *Challenging Behaviour: Analysis and Intervention in People with Severe Intellectual Disabilities*. Cambridge University Press: Cambridge

Emmer, E.T., & Hickman, J. (1992). Teacher Efficacy in Classroom Management and Discipline. *Educational and Psychological Measurement*. **51**(3), 755-765.

Evans, J., Harden, A., Thomas, J., & Benefield, P. (2003). *Support for pupils with emotional and behavioural difficulties (EBD) in mainstream primary classrooms: a systematic review of the effectiveness of interventions*. London: EPPI-Centre, Social Sciences Research Unit, Institute of Education, University of London.

Egyed, C. J., & Short, R. J. (2006). Teacher self-efficacy, burnout, experience, and decision to refer a disruptive student. *School Psychology International*. **27**(4), 462-474.

Federici, R.A., & Skaalvik, E.M. (2012). Principal self-efficacy: relations with burnout, job satisfaction and motivation to quit. *Social psychology of education*. **15**(3), 295-320.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. 4th Ed. Sage Pulications Inc: London.

Ford, T., Edwards, V., Sharkey, S., Ukoumunne, O.C., Byford, S., Norwich, B., & Logan, S. (2012). Supporting teachers and children in schools: the effectiveness and cost-effectiveness of the incredible years teacher classroom management programme in primary school children: a cluster randomised controlled trial, with parallel economic and process evaluations. *BMC Public Health*. **12**, Article 719.

Fox, M. (2009). Working with systems and thinking systematically – disentangling the crossed wires. *Educational Psychology in Practice: theory, research and practice in educational psychology*. **25**(3), 247-258

Gable, R.A, & Hendrickson, J.M. (2000). Strategies for maintaining Positive behavior change Stemming from Functional Behavioral Assessment in Schools. *Education and Treatment of Children*. **23**(3), 286-297.

Giallo, R., & Little, E. (2003). Classroom Behaviour Problems: The Relationship between Preparedness, Classroom Experiences, and Self-efficacy in Graduate and Student Teachers. *Australian Journal of Educational & Developmental Psychology*. **3**(2), 21-34.

Gibbs, S., & Powell, B. (2012). Teacher efficacy and pupil behaviour: The structure of teachers' individual and collective beliefs and their relationship with numbers of pupils excluded from school. *British Journal of Educational Psychology*. **82**(4), 564-584.

Gibson, S., & Dembo, M.H. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology*. **76**(4), 569-582.

Gisev, N., Bell, J.S., & Chen, T.F. (2013). Inter-rater agreement and inter-rater reliability: Key concepts, approaches, and applications. *Administrative Pharmacy*. **9**(3), 330-338.

Goh, A.E., & Bambara, L.M. (2010). Individualized Positive Behavior Support in School Settings: A Meta-Analysis. *Remedial and Special Education*. **33**(5),271-286.

Gorard, S. & Taylor, S. (2004). *Combining Methods in Educational and Social Research*. Open University Press: McGraw – Hill Education.

Gough D (2007) Weight of evidence: a framework for the appraisal of the quality and relevance of evidence. In J. Furlong, A. Oancea (Eds.) Applied and Practice-based Research. *Special Edition of Research Papers in Education*. **22** (2), 213-228

Hall, E., Hall, C., & Abaci, R. (1977). The effects of human relations training on reported teacher stress, pupil control ideology and locus of control. *British Journal of Education*. **67**(2), 483-496.

Hargreaves, D. J., Purves, R. M., & Graham, F. (2007). Developing identities and attitudes in musicians and classroom music teachers. *British Journal of Educational Psychology*. **77**(3), 665-682.

Hastings, RP., & Bham, M.S (2003). The Relationship between Student Behaviour Patterns and Teacher Burnout. *School Psychology International.* **24**(1), 115-129.

Hastings, R.P., & Brown, T. (2002). Behavioural knowledge, causal beliefs and self-efficacy as predictors of special educators' emotional reactions to challenging behaviours. *Journal of Intellectual Disability Research*. **46**(2), 144-150.

Hayden, T. (2014). To what extent is behaviour a problem in English schools? Exploring the scale and prevalence of deficits in classroom climate. *Review of Education.* **2**(1), 31-64.

Hemmeter, M.L., Fox, L., Jack, S., & Broyles, L. (2007). A Program-Wide Model of Positive Behavior Support in Early Childhood Settings. *Journal of Early Intervention*. **29**(4), 337-355.

Higgens, J.P.T., & Altman, D.G. (editors). (2008). Chapter 8: Assessing risk of bias included in studies. In Higgens, J.P.T., & Green, S. (editors). (2008). Cochrane Handbook for Systematic Reviews of Interventions. John Wiley & Sons: Chichester (UK)

Higgins, H., & Gulliford, A. (2014). Understanding teaching assistant self-efficacy in role and in training: its susceptibility to influence. Educational Psychology in Practice: theory, research and practice in educational psychology. 30(2), 120-138.

Hong, J.Y. (2012). Why do some beginning teachers leave the school, and others stay? Understanding teacher resilience through psychological lenses. *Teachers and Teaching*. **18**(4), 417-440.

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.

Howitt, D., & Cramer, D. (2008). *Introduction to SPSS in Psychology For Version 16 and earlier*. Fourth Edition. Pearson Education Ltd: London.

Jennet, H.K., Harris, S.L., & Mesibov, G.B. (2003). Commitment to Philosophy, Teacher Efficacy, and Burnout Among Teachers of Children with Autism. *Journal of Autism and Developmental Disorders*. **33**(6)**,** 583-596.

Johnson, S.M., & White, G. (1971). Self-Observation as an Agent of Behavioral Change. *Behavior Therapy.* **2**, 488-497.

Joseph, S., & Murphy, D. (2012). Person-Centered Approach, Positive Psychology, and Relational Helping: Building Bridges. *Journal of Humanistic Psychology*. **53**(1), 26-51.

Kaspereen, D. Relaxation Intervention for Stress Reduction Among Teachers and Staff. *International Journal of Stress Management.* **19**(3), 238-250.

Kazdin, A.E. (1984). Statistical Analysis for Single-case Experimental Designs (Chapter 9). In Barlow, D.H., & Hersen, M. (1984). *Single Case Experimental Designs. Strategies for studying behaviour change (Second Edition*). Pergamon Press Inc: Oxford.

Kazdin, A. E. (2003). *Research Design in Clinical Psychology*. (4th Ed.). Allyn & Bacon: Boston.

Kim, J.S. (2008). Examining the Effectiveness of Solution-Focused Brief Therapy. *Research on Social Work Practice*. **18**(2), 107-116.

Kim, J.S., & Franklin, C. (2009). Solution-focused brief therapy in schools: A review of the outcome literature. *Children and Youth Services Review*. **31**, 464-470.

Klingner, J. K., Ahwee, S., & Pilonieta, P. (2003). Barriers and facilitators in scaling up research-based practices. *Exceptional Children.* **69**(4), 411-429.

Kokkinos, C. M. (2007). Job stressors, personality and burnout in primary school teachers. *The British Journal of Educational Psychology.* **77**(1), 229–243.

Kratochwill, T., Hitchcock, J., Horner, R., Levin, J., Odom, S., Rindskopf, D., & Shadish, W. (2010). *Single-Case Designs Technical Documentation.* Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Kurt, T., Duyar, I., & Calik, T. (2012). Are we legitimate yet? A closer look at the casual relationship mechanisms among principal leadership, teacher self-efficacy and collective efficacy. *Journal of Management Development*. **31**(1), 71-86.

Lang, R., Sigafoos, J., Lanconi, G., Didden, R., & Rispoli, M. (2010). Influence of assessment setting on the results of Functional analyses of problem behaviour. *Journal of Applied Behavioural Analysis.* **43**(3), 565-567.

Lee, B., Cawthon, S., & Dawson, S. (2013). Elementary and secondary teacher self-efficacy for the teaching and pedagogical conceptual change in a drama-based professional development program. *Teaching and Teacher Education*. **30**(31), 84-98.

Leyser, Y., Zeiger, T & Romi, S. (2011). Changes in Self-efficacy of Prospective Special and General Education Teachers: Implication for inclusive education. *International Journal of Disability, Development and Education.* **58**(3), 241-255.

Liaw, E. (2009). Teacher efficacy of pre-service teachers in Taiwan: The influence of classroom teaching and group discussions. *Teaching and Teacher Education*. **25**(1), 176-180.

Lokke, C., Gersch, I., M'gadzah, H., & Frederickson, N. (1997). The Resurrection of Psychometrics. *Educational Psychology in Practice: theory, research and practice in educational psychology*. **12**(4), 222-223.

Loveridge, J. (2010). Involving Children and Young People in Research in Educational Settings. *Report to the Ministry of Education*.

Mackenzie, N., & Knipe, S. (2006). Research dilemmas: Paradigms, methods and methodology. *Issues in Educational Research*. **16**(2), 193-205.

Macleod, F., Macmillan, P., and Norwich, B. (2007) Giving Psychology Away: Helping Pupils at Risk of Reading Failure By Means of a 288 Self-Voice Feedback Programme. *School Psychology International.* **28**(5), 555-573.

Maletzky, B.M. (1974). Behavior Recording as Treatment: A Brief Note. *Behaviour Therapy*. **5**, 107-111.

Manolov, R., & Solanas, A. (2008). Comparing N = 1 Effect Size Indices in Presence of Autocorrelation. *Behavior Modification.* **32**(6), 860-875.

Maruyama, G., & Deno, S. (1992). Researc*h in Educational Settings*. Sage Publications: California

Maslach, C., Jackson, S.E., & Leiter, M.P. (1996). *The Maslach Burnout Inventory* (3rd ed). Consulting Psychologists Press: Palo Alto, CA.

Maslach, C. (2003). Job Burnout: New Directions in Research and Intervention. *Current Directions in Psychological Science*. **12**(5), 189-192.

Maslach, C., & Jackson, S.E. (1981). The measurement of experienced burnout. *Journal of organisational behavior.***2**(2), 99-113.

Matyas, T.A., & Greenwood, K.M. (1990). Visual analysis of single-case time series: effects of variability, serial dependence, and magnitude of intervention effects. *Journal of Applied Behavior Analysis*. **23**(3), 341-351.

McIntosh, K., Brown, J.A., & Borgmeister, C.J. (2008). Validity of Functional Behavior Assessment Within a Response to Intervention Framework: Evidence**.** *Assessment for Effective Intervention*. **34**(1), 6-14.

McNamara, E. (1979). The use of Self-Recording in Behaviour Modification in a Secondary School. *Behavioural Psychotherapy*. **7**(3), 57-66.

Merrett, F., & Wheldall, K (1987). The Behavioural Approach to Teaching. In Wheldall, K. (1987). *The Behaviourist in the Classroom*. Allen & Unwin: London.

Merrett, F., & Wheldall, K. (1993). How Do Teachers Learn to Manage Classroom Behaviour? A study of teachers' opinions about their initial training with special reference to classroom behaviour management. *Educational Studies.* **19**(1), 91-106.

Mertens, D.M. (2005). *Research and Evaluation in Education and Psychology*. 2nd Ed. Sage Publications Inc: London.

Miller, A. (2008b). How ethical are behavioural approaches to classroom management? In Frederickson, N., Miller, A., & Cline, T. (2008). *Educational Psychology. Topics in Applied Psychology*. Hodder Education: London

Miller, G., & de Shazer, S. (2000). Emotions in solution-focused therapy: A re-examination. *Family Process.* **39**(1), 5−23.

Mitchell, G., & Hastings, R.P. (2001). Coping, Burnout, and Emotion in Staff Working in Community Services for People with Challenging Behaviors. *American Journal on Mental Retardation*. **106**(5), 448-459.

Moraco, J.C., & McFadden, H. (1982). The Counselor's Role in Reducing Teacher Stress. *The Personnel and Guidance Journal*. May, 549 – 554

Moreno, G. (2010). No need to count to ten: advocating for the early implementation of the Functional Behavioural Assessment in addressing challenging behaviours. *Emotional and Behavioural Difficulties*. **15**(1), 15-22.

Mossman, D.A., Hastings, R.P., & Brown, T. (2002). Mediators' emotional responses to self-injurious behaviour: an experimental study. *American Journal of Mental Retardation.* **107**(4), 252-260.

Müller, K, F., Briel, M., D'Amariom A., Kleijnen, J., Marusic, A., Wager, E., Antes, G., von Elm, E., Lang, B., Motschall, E., Gloy, V., Schwarzer, G., Altman, D., Meerpohl, J.J., Bassler, D. (2013). Defining publication bias: protocol for systematic review of highly cited articles and proposal for a new framework. *Systematic Reviews.* **2:**34.

Newman-Carlson, D., & Horne, A.M. (2004). Bully Busters: A Psychoeducational Intervention for Reducing Bullying Behaviour in Middle School Students. *Journal of Counselling & Development.* **82**(3), 259 – 271.

National Foundation for Educational Research (NFER). (2012). Teacher Voice Omnibus February 2012 survey: Pupil behaviour. *Department for Education.*

OFSTED. (2010). *The Annual Report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2009/10*. Crown Copyright

Officer, B. (2012). No correlation between middle school teachers' training in adolescent mental health and teachers' self-efficacy. *Dissertation Abstracts International: Section B: The Sciences and Engineering*. **73**(2-B), 872.

Ozga, J. (2000). *Policy Research in Educational Settings contested terrain*. Open University Press: Buckingham.

Pallant, J. (2013). *SPSS Survival Manual.*5th Ed. McGraw-Hill: Berkshire.

Palmer, D. (2010). Sources of Efficacy Information in an Inservice Program for Elementary Teachers. *Science Education.* 577- 601.

Pameijer, N. (2006). Towards needs-based assessment: Bridging the gap between assessment and practice. *Educational & Child Psychology*. **23**(3), 12- 24.

Parker, R.I., Vannest, J., & Davis, J.L. (2011). Effect Size in Single-Case Research: A Review of Nine NonoverlapTechniques. *Behavior Modification.* **35**(4), 303-322.

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation.* Sage Publications: London.

Pelligrini, D.W. (2009). Applied systematic theory and educational psychology: can the twain ever meet? *Educational Psychology in Practice: theory, research and practice in educational psychology*. **25**(3), 271-286.

Penrose, A., Perry, C., Ball, I. (2007). Emotional intelligence and teacher self efficacy: The contribution of teacher status and length of experience. *Issues in Educational Research*. **17**(1), 107-126.

Porter, L. (2007). *Behaviour in Schools. Theory and Practice for Teachers.* 2nd Ed. Open University Press: Berkshire

Prochaska, J. O., & Diclemente, C. C. (1998) *Comments, criteria and creating better models: in response to Davidson, in: W. Miller & N. Heather (Eds) Treating addictive behaviours (2nd Ed*) .Plenum Press: New York.

Purdey, B. (2013). Occupant stimulus response workplace productivity and the vexed question of measurement. *Workplace Productivity*. **31**(11-12), 505-520.

Qiao, H., & Schaufeli, W.B. (2011). The Convergent Validity of Four Burnout Measures in a Chinese Sample: A Confirmatory Factor – Analytic Approach. Applied Psychology: *An International Review*. **60**(1), 87-111.

Rae, T., & Daly, S. (2008). *Controlling Anger. A solution focused approach for children*. Optimus Education: London

Reese, R.M., Sherman, J.A., & Sheldon, J. (1984). Reducing Agitated-Disruptive Behavior of Mentally Retarded Residents of Community Group Homes: The Role of Self-Recording and Peer-Prompted Self-Recording. *Analysis and Intervention in Developmental Disabilities.* **4**(2), 91-107.

Revital, S.S. (2009). Dealing with school violence: The effect of school violence prevention training on teachers' perceived self-efficacy in dealing with violent events. *Teaching and Teacher Education*. **25**(8), 1061-1066.

Robertson, C., & Dunsmuir, S. (2013). Teacher stress and pupil behaviour explored through a rational-emotive behaviour therapy framework. *Educational Psychology*. **33**(2), 215-232.

Robson, C. (2011). *Real World Research.* 3rd ed. John Wiley and Sons Ltd: Cornwall

Ross, S, G. (2012). *Measuring Response to Intervention: Comparing Effect Size Calculation Techniques for Single-Case Design Analysis.* A Dissertation submitted to the Graduate Faculty of North Carolina State Univertsity.

Royal College of Psychiatrists, British Psychological Society and Royal College of Speech and Language Therapists. (2007). *Challenging Behaviour: a unified approach*. Royal College of Psychiatrists.

Shabani, D.B., Wilder, D.A., & Flood, W.A. (2001). Reducing Stereotypic behaviour through discrimination training, differential reinforcement of other behaviour and self-monitoring. *Behavioural Interventions.* **16**, 279-286.

Shechtman, Z., Levy, M., & Leichtentritt, J. (2005). Impact of life skills training on teachers' perceived environment and self-efficacy. *Journal of Educational Research.* **98**(3), 144-154.

Shirom, A., & Melamed, S. (2006). A Comparison of the Construct Validity of Two Burnout Measures in Two Groups of Professionals. *International Journal of Stress Management*. **13(**2), 176-200.

Skinner, B.F. (1971). *Beyond Freedom & Dignity.* Hackett Publishing Company, Inc.: Indiana.

Steer, A. (2009). *Learning Behaviour: LESSONS LEARNED. A review of behaviour standards and practices in our schools*. Institute of Education : London.

Stoiber, K.C., & Gettinger, M. (2011). Functional assessment and positive support strategies for promoting resilience: effects on teachers and high risk children. *Psychology in the schools.* **48**(7), 686-706.

Sugai, G., Horner, R.H., Dunlap, G., Hieneman, M., Lewis, T.J., Nelson, C.M., Scott, T., Liaupsin, C., Sailor, W., Turnbull, A.P., Rutherford Turnball III, H.,

Wickham, D., Wilcox, B., & Reuf, M. (2000). Applying Positive Behaviour Support and Functional Behavioral Assessment in Schools. *Journal of Positive Behavior Interventions*. 2, 131- 146.

Teaching Agency. (2012). *Improving Teacher Training for Behaviour.* Crown Copyright.

Telljohann, S.K., Everett, S.A., Durgin, J., & Price, J.H. (1996). Effects of an inservice workshop on the health teaching self-efficacy of elementary school teachers. *Journal of School Health*. **66**(7), 261-265.

Tschannen-Moran, M., Woolfok Hoy, A., & Hoy, W. (1998). Teacher Efficacy: Its meaning and measure. *Review of Educational Research*. **68**(2), 202-248.

Ware, H., & Kitsantas, A. (2007). Teacher and Collective Efficacy Beliefs as Predictors of Professional Commitment. *Journal of Educational Research.* **100**(5), 303-310.

Watkins, C., & Wagner, P. (2000). *Improving School Behaviour.* Paul Chapman Publishing Ltd: London.

Weisberg, J., & Sagie, A. (1999). Teachers' Physical, Mental, and Emotional Burnout: Impact on Intention to Quit. *The Journal of Psychology: Interdisciplinary and Applied Psychology*. **133**(3), 333-339.

Winnet, R.A., & Winkler, R.C. (1972). Current behavior modification in the classroom: Be still, be quiet, be docile. *Journal of Applied Behavioral Analysis.* **5**(4), 499-504

Wolters, C.A., & Dougherty, S.G. (2007). Goal Structures and Teachers' Sense of Efficacy: Their Relation and Association to Teaching Experience and Academic Level. *Journal of Educational Psychology*. **99**(1), 181 – 193.

Wood, B.K., Cho Blair, K-S., & Ferro, J.B. (2009). Young Children with Challenging Behaviour: Function-Based Assessment and Intervention. *Topics in Early Childhood Special Education.* **29**(2), 68 – 80.

Secondary References

Farber, B. A. (1991). Crisis in education: Stress and burnout in the american teacher. In Jennet, H.K., Harris, S.L., & Mesibov, G.B. (2003). Commitment to Philosophy, Teacher Efficacy, and Burnout Among Teachers of Children with Autism. *Journal of Autism and Developmental Disorders*. **33**(6)**,** 583-596**.**

Salzberg, (1972). In Pereira, J.A., & Winton, A.S.W. Teaching and remediation of mathematics: A review of behavioral research. *Journal of Behavioral Education*. **1**(1),5-36.

Websites and Electronic resources

DEPARTMENT FOR EDUCATION, (2012c). Press Release: Trainee teachers to get a better grip on managing behaviour. [online]. Department for Education, 2nd July 2012. [Viewed 29th July 2014]. Available from: https://www.gov.uk/government/news/trainee-teachers-to-get-a-better-grip-on-managing-behaviour--2

DOMOKOS, J. Illegal school exclusions: How pupils are slipping through the net. [online]. The Guardian Education, 15th November 2012. [Viewed 29th July 2014]. Available online at: http://www.theguardian.com/education/2012/nov/15/illegal-school-exclusions

Shirom, A., & Melamed, S. (2005). *Shirom-Melamed Burnout Measure*. ARIE SHIROM. Proff Arie Shirom, [online]. Tel Aviv. Tel Aviv University 2012. [Viewed 29th July 2014]. Available from: http://www.shirom.org/arie/index.html

SELLGREN, K. Disruptive behaviour rising, teachers say. [online]. BBC News, 23rd March. 2013 [Viewed 22nd July 2013]. Available from: http://www.bbc.co.uk/news/education-21895705.

OFFICE FOR NATIONAL STATISTICS. Office for National Statistics [online]. London ONS, 2011. [Viewed 29th July 2014]. Available via http://www.ons.gov.uk/ons/guide-method/census/2011/index.html

TEACHFIRST. TeachFirst [online]. London TeachFirst, 2014. [Viewed 28th July 2014]. Available from: http://www.teachfirst.org.uk/

VASAGER, J. *Children's behaviour at school deteriorating, say teachers.* The Guardian, 18th April 2011. [Viewed 29th July 2014]. Available from: http://www.guardian.co.uk/education/2011/apr/18/childrens-behaviour-school-teachers

9 Appendices

Appendix 1: Systematic Review Search Criteria, Results and Exclusions

Psych INFO

1) "teacher" - exp Junior High School Teachers/ or exp High School Teachers/ or exp Special Education Teachers/ or exp Student Teachers/ or exp Elementary School Teachers/ or exp Middle School Teachers/ = **15598**
2) "self-efficacy" - exp Self-efficacy/ = **13414**
3) **"Burnout" -** burnout.mp. or exp Occupational Stress = **16824**
4) **"Training" -** exp Social Skills Training/ or exp On the Job Training/ or training.mp. or exp Training

Criteria - published peer reviewed article, school age children, teaching staff in schools

- Full article (no dissertation abstracts)
- Training must relate to teaching (not use of technology etc.)
- Self-efficacy / training - must be influenced / measured in relation to training given

SEARCH 1: 1 AND 2 AND 3 AND 4

2 results – one removed as abstract only available

One full text remaining:

Jennet et al (2003)

SEARCH 2: 1 AND 2 AND 4

32 results

22 excluded based on reading of abstracts

- removed due to abstract only availability, outcomes not being self-efficacy/teacher burnout.

Resulted in 10 studies, removed one duplicate

Nine studies to read full text:

Britten & Lai (1998)

Chan, (2005) – removed as no intervention (Q'nnaire based study correlating measures)

Clerici (2008) – disregarded as no intervention put in place

Palmer (2010)

Lee et al (2013)

Newman-Carlson & Horne (2004)

Meet criteria according to title and/or abstract but full text not available:

Yona, L., Tali, Z., Shlomo, R. (2011). Changes in self-efficacy of prospective special and general education teachers: Implication for inclusive education. International Journal of Disability, Development and Education. Changes in self-efficacy of prospective special and general education teachers: *Implication for inclusive education*. **58**(3), 241-255.

Milson, A.J. (2003). Teachers' sense of efficacy for the formation of students' character. *Journal of Research in Character Education*. **1**(2), 89-106

Watkins, D. (2000). Hong Kong student teachers' personal construction of teaching efficacy. *Educational Psychology.* **20**(2), 213-235.

SEARCH 3: 1 AND 3 AND 4

23 studies

Excluded based on their abstracts:

3 removed – no English text available

1 removed – duplicate

2 removed – adult military training/adult education

12 removed – dissertation abstract only

2 removed – self-efficacy or burnout (or clearly related concepts) not an outcome

5 Studies to read full text:

Hall et al (1997)

Maracco & Hope (1982)

Meet criteria according to title and/or abstract but full text not available:

Seidman, S. A., Zager, J. (1992). Teacher stress workshops. *Work & Stress.* **6**(1), 85-87.

Cox, T., Boot, N., Cox, S., Harrison, S. (1988). Stress in schools: An organizational perspective. *Work & Stress.* **2**(4), 353-362.

Woodhouse, D. A., Hall, E., Wooster, A. D. (1985). Taking control of stress in teaching. *British Journal of Educational Psychology*. **55**(2), 119-123.

WEB OF KNOWLEDGE

1) Search Topic: teacher and self-efficacy and burnout and training

7 results

Exclusions through abstracts/titles

4 – no intervention (interview retrospective based/q'nnaire based)

1 – adult education

1 duplicate – Jennet et al (2003)

For full text to be read excluded as not available in English:

Wudy, D.T., Jerusalem, M. (2011). Changes in Teachers' Self-Efficacy and Experiences of Stress. *Psychologie in erzeihung und unterricht*. **58**(4), 254-267.

2) Search Topic: teacher and self-efficacy and training

223 results

Refine by "school", "article" (document type),

110 results – Read titles and abstracts:

3 removed – not available in English text/unavailable

41 removed – no intervention

43 - removed – teacher self-efficacy/burnout - not an outcome (other participants are focus e.g students/parents etc./ or other constructs are focus that are related to self-efficacy).

4 removed – adult education/not an educational setting

9 – self-efficacy and/or burnout as a independent rather than dependent variable

1 – duplicate

Articles relevant after exclusion by abstract and title (9):

Ford et al (2012).

Stoiber, & Gettinger (2011).

Revital (2009).

Liaw (2009).

Hargreaves et al (2007).

Shechtman et al (2005).

Telljohann et al (1996).

Meet criteria but not available:

Kantor, G.K., Caudill, B.D., Underleider, S. (1992). Project Impact – teaching the teachers to intervene in student substance abuse problems. *Journal of alcohol and drug education*. **38**(1), 11-29.

Tanaka, S. (2006). Teacher-training students' preparation and self-efficacy regarding skills of teaching preschool children : Observing peer models and modelling. *Japanese Journal of Educational Psychology*. **54**(3), 408-419. -

3) Search Topic: teacher and burnout and training

70 articles

Refine by “school”, “article” (document type),

36 results – Read titles and abstracts:

2 removed – not available in English text/unavailable

23 removed – no intervention

3 - removed – teacher self-efficacy/burnout - not an outcome (other participants are focus e.g students/parents etc./ or other constructs are focus that are related to self-efficacy).

4 removed – adult education/not an educational setting

2 removed - self-efficacy and/or burnout as a independent rather than dependent variable

Articles relevant after exclusion by abstract and title (2):

Kaspereen (2012).

Meets criteria but not available:

Tyson, O., Roberts, C. M., Kane, R. (2009). Can Implementation of a Resilience Program for Primary School Children Enhance the Mental Health of Teachers? *Australian Journal of Guidance and Counselling*. **19**(2), 116-130.

Appendix 2: Tables of Studies for Systematic Review

1. Jennet, H.K., Harris, S.L., Mesibov, G.B. (2003). Commitment to Philosophy, Teacher Efficacy, and Burnout Among Teachers of Children with Autism. *Journal of Autism and Developmental Disorders.* **33**(6)**, 583-596.**

Participants	34 teachers using ABA, 30 using TEACCH (with young people with ASC)
Intervention	Standard ABA or TEACCH training
Outcomes	Autism treatment philosophy Questionnaire (to distinguish between approaches), Maslach burnout inventory, Teacher Efficacy Scale.
Design	Comparison, questionnaire
Comparisons	TEACCH and ABA teachers
Results	TEACHH teachers – had significantly higher TEACCH score than ABA, ABA significantly higher ABA score than TEACHH – shared dimension – ABA group significantly higher association with this. Groups – no significant difference between group burnout or self-efficacy (overall high self-efficacy, low burnout). Both groups low depersonalisation Commitment score – significantly correlated (positively) with personal self-efficacy scores ($p<0.05$, $p<0.001$) Commitment score (for ABA group only) – significant positive correlation with general self-efficacy scores ($p< .001$) TEACCH – higher commitment score = lower emotional exhaustion and higher personal accomplishment ($p<0.05$) ABA – approached significance – but not ($p= 0.1$)
Outcome Measures	Philosophical commitment, teacher self-efficacy scale, Maslach Burnout Inventory
Outcome sizes	
Blinding of Assessors	
Participant Selection	Teachers identified based on orientation of programme they worked on. Potential participants sent questionnaires through the post.

Main Investigation Question	Teachers who identify themselves with specific teaching (ABA/TEACCH) would be more committed to the philosophy of that teaching. Teachers who have high commitment to teaching philosophy – high self-efficacy and low burnout
Potential Limitations	There were significant contributions of individual variables on scores (although not significant)
Reliability	

2. Britten, P., Lai, M.K. (1998). Structural Analysis of the Relationships among Elementary Teachers' Training, Self-Efficacy, and Time Spent Teaching Nutrition. *Journal of Nutrition Education.* **30**, 218 - 224.

Participants	324 Elementary teachers
Intervention	Theoretical – model – structural analysis
Outcomes	Nutrition self-efficacy, nutrition training, time spent teaching nutrition
Design	Questionnaire Exploratory factor analysis of previous research to determine which questions are relevant to the Research Question 4 Q's – nutrition training 2 Q's – self-efficacy 2 Q's – Time spent teaching nutrition 24Q's – nutrition knowledge 4 Q's - Belief it is important to teach nutrition Covariance analysis of linear structural equation – to test the row linear models to see which best fits with the data
Comparisons	Compared to models to alternative model which was more flexible.
Results	Model 1 – path = self-efficacy to time spent training to self-efficacy and knowledge to self-efficacy = ($p<0.05$) Model 2 – path = self-efficacy to time spent, training to time spent, knowledge to self-efficacy = ($p<0.05$) Primary model and alternative model (allowing for both

	direct and indirect impact on teaching) fitted data better than second model Primary and alternative – difference = 0.81, secondary and more flexible – $p<0.01$. Belief that nutrition teaching is important – no correlation to time spent teaching. Training and knowledge - no correlation to time spent teaching. Nutrition knowledge – significant correlation with self-efficacy In service training – minimal if any impact on time spent teaching.
Outcome Measures	Self-efficacy, nutrition teaching time, training
Outcome sizes	
Blinding of Assessors	N/A
Ppt Selection	Used from a previous study – 6 schools randomly selected from state – head teachers asked to ask staff members to complete questionnaire
Main Investigation Question	1) Teachers with more nutrition training will have increased self-efficacy for teaching nutrition and therefore increased time spent teaching nutrition. 2) Training and self-efficacy will directly and independently influence time spent teaching nutrition.
Potential Limitations	No measures of quality of teaching. No student outcome measure.
Reliability	Self-report, retrospective analysis Measures – not designed for this purpose (e.g. only 2 self-efficacy measures)

3. Newman-Carlson, D., Horne, A.M. (2004). Bully Busters: A Psychoeducational Intervention for Reducing Bullying Behaviour in Middle School Students. *Journal of Counseling & Development.* **82**, 259 – 271.

Participants	6^{th}, 7^{th} and 8^{th} grade middle school teachers – public

	school – southeastern United state
Intervention	Bully prevention program
Outcomes	Teachers knowledge, use of bullying intervention skills, teacher self-efficacy, student classroom bullying behaviour.
Design	3 training sessions and a support team. Quasi experimental pre-test post-test control group design.
Comparisons	
Results	Q 1 – (knowledge) – All ANCOVAs = significant treatment group – significantly higher knowledge (reject null) Q2 – All ANCOVAs significant (reject null) Q3 – Significantly higher teacher self-efficacy for 5 of the 7 typologies (average, disruptive behaviour disorder, learning disorder, severe psychopharmacology, mildly disruptive) – Not physical complaints / worry or average. Also not on general self-efficacy (reject null) Q4 – ANCOVA = significant (reject null)
Outcome Measures	Knowledge (Teacher Inventory of Skills and Knowledge), self-efficacy (teacher efficacy scale), efficacy and attribution (Teacher efficacy and attribution Measure), use of intervention (Osiris School Administration Activity Tracker).
Outcome sizes	All significant $p<0.01$ or $p<0.05$
Blinding of Assessors	No
Ppt Selection	15 teachers – 42 teachers offered the training – received continuing education credit. 15 controls (those teachers who declined to participate in training)
Main Investigation Question	1 – Does a psychoeducational intervention for middle school teachers affect teachers' knowledge of bullying intervention skills? 2- Does a psychoeducational intervention for middle school teachers affect teachers use of bullying interventions? 3 – Does a psychoeducational intervention for middle

	school teachers affect teachers' self-efficacy? 4 – Does a psychoeducational intervention for middle school teachers have an effect on the number of student disciplinary referrals?
Potential Limitations	Author – multiple roles
Reliability	Across instructors?

Removed on Full Text Analysis:

Chan, D.W. (2005). Counselling values and their relationships with self-efficacy among Chinese secondary school teachers in Hong Kong. *Counselling Psychology Quarterly.* **18**(3), 183-192. – No training/intervention evaluated and self-efficacy is independent not a dependent variable.

Clerici, R. (2008). Knowledge and attitudes of future schoolteachers in the scientific-mathematical sphere: some evidences for Italy. *Educational Studies.* **34**(4), 277-287. – No intervention used

4. Palmer, D. (2010). Sources of Efficacy Information in an Inservice Program for Elementary Teachers. *Science Education.* 577- 601.

Participants	12 teachers from small elementary schools in southeastern Australia (Grade 3-4), minimum 4 years teaching – ages 20-45
Intervention	Designed to give teachers cognitive mastery, enactive mastery, modelling and verbal persuasion. (training spread over 6 weeks) Workshop phase, observation phase, teaching phase – aimed to enhance self-efficacy for teaching science
Outcomes	Self-efficacy
Design	Pre- intervention, immediate post, delayed group (2 years)
Comparisons	Pre-post
Results	Main increases in self-efficacy were mainly due to cognitive mastery (effect size - 1.24) and in situ feedback (effect size – 1.69). Self-efficacy – pre-test significantly lower than post and delayed (p=0.002)

	No significant difference between immediate post and delayed Workshop most commonly identified as most useful
Outcome Measures	10 occasions Survey – science teaching efficacy belief instrument Audiotaped interviews (semi-structured) Questionnaires – closed and open question
Outcome sizes	
Blinding of Assessors	
Ppt Selection	
Main Investigation Question	1 – What are the effects of cognitive mastery and enactive mastery as sources of efficacy for elementary teachers? 2- What is the effect of vicarious experience provided for elementary teachers in their own classroom? 3 – What is the effect of repetitious familiarity in relations to fear and stress among elementary teachers? 4 – What is the comparative effectiveness of cognitive mastery, enactive mastery, in situ modelling, in situ feedback and repetitious familiarity as sources of efficacy information for elementary teachers?
Potential Limitations	Small sample
Reliability	Data sources triangulated

5. Lee, B., Cawthon, S., Dawson, S. (2013). Elementary and secondary teacher self-efficacy for the teaching and pedagogical conceptual change in a drama-based professional development program. *Teaching and Teacher Education.* **30**, 84-98.

Participants	12 elementary school teachers, 18 secondary school teachers
Intervention	TAT professional development programme (drama based teaching techniques to use in the classroom)
Outcomes	Self-efficacy

Design	Mixed methods
Comparisons	Pre- and post- secondary and elementary
Results	Pre self-efficacy scores in classroom management and student engagement significantly higher for elementary ($p<0.05$) (not significantly different for instructional strategies self-efficacy scale) Pre- and post- – no significant difference for elementary Post- – elementary conceptual change altered significantly ($p<0.001$) Secondary teachers – as student engagement self-efficacy increased conceptual change increased
Outcome Measures	Evaluation of training, Self-efficacy scale, lesson plan evaluations
Outcome sizes	
Blinding of Assessors	No
Ppt Selection	Volunteers
Main Investigation Question	Elementary teachers have higher initial self-efficacy than secondary teachers Elementary teachers experience greater conceptual change than secondary school teachers Teachers with higher self-efficacy have greater shifts in conceptual change than teachers with lower self-efficacy
Potential Limitations	High drop-out of participants for final self-efficacy ratings as used on-line method for this data collection only (14 participants)
Reliability	Account for number of years of teaching experience in analysis – no effect found on any of the measures.

6. Hall, E., Hall, C., Abaci, R. (1977). The effects of human relations training on reported teacher stress, pupil control ideology and locus of control. *British Journal of Education.* **67**, 483-496.

Participants	42 experienced teachers, 42 control
Intervention	Two years Masters programme in human relations (experiential learning – practical learning situations). 3

	hours a week for 5 consecutive 10 week terms
Outcomes	Post training – reduction in reported stress, increased humanistic orientation towards classroom management, increased internal locus of control.
Design	RCT
Comparisons	Pre- and post- training, and control group
Results	Experimental group: reduction in emotional exhaustion ($p<.01$) increase in sense of personal accomplishment ($p<.05$) reduction in depersonalisation (not significant) Significant shift to humanistic ideology for experimental but not control (group difference $p<.01$) Interview results suggest : Experimental group stress reduced – explained as more implementation of strategies to reduce stress Movement towards greater internal sense of control
Outcome Measures	Maslach burnout inventory, Pupil control ideology form, semi-structured interviews
Outcome sizes	Unknown
Blinding of Assessors	NO
Ppt Selection	42 experienced volunteer teachers. Volunteers suggested a colleague who is similar to them for control group
Main Investigation Question	As a result of 2 year Masters programme in human relations for teachers does: Reported stress reduce? Attitudes to classroom management become more humanistic? Participants feel an increased sense of control over their lives? Work behaviour change?

Potential Limitations	Self-report, 2 year course – generalisability
Reliability	Interviewers knew aims of research

Removed on Full Text Analysis:

Moraco, J.C., McFadden, H. (1982). The Counselor's Role in Reducing Teacher Stress. *The Personnel and Guidance Journal.* May, 549 – 554 - Removed as no intervention evaluated – use theory to guide a premise for an intervention.

Ford, T., Edwards, V., Sharkey, S., Ukoumunne, O.C., Byford, S., Norwich, B., & Logan, S. (2012). Supporting teachers and children in schools: the effectiveness and cost-effectiveness of the incredible years teacher classroom management programme in primary school children: a cluster randomised controlled trial, with parallel economic and process evaluations. *BMC Public Health.* **12**, Article 719. – Results not yet published.

7. Stoiber, K.C., Gettinger, M. (2011). Functional assessment and positive support strategies for promoting resilience: effects on teachers and high risk children. *Psychology in the schools.* **48**(7), 686-706.

Participants	70 teachers (pre-kindergarten, kindergarten and first grade) 90 students (2 nominated from each class by teacher)
Intervention	Professional development training – FBA and positive behaviour support
Outcomes	Teacher resilience, self-efficacy and use of FBA procedures Student behaviour
Design	RCT pre- and post- (grouped by district and then randomly assigned to condition)
Comparisons	35 experimental condition to 35 control
Results	Significantly higher ratings of competence and self-efficacy for experimental group ($p<0.01$) Significantly higher ratings from observations of utilization of skills for experimental group ($p<0.001$)

	(But significantly higher for target children than general children) Significant improvements for target children compared to control and general ($p<0.01$) Significant improvements between general and control ($p<0.05$)
Outcome Measures	Staff ratings: Competency self-ratings Accommodating children with challenging behaviour questionnaire Observer Rating of Eco behavioural Variables Scale Student measures: Social competence performance checklist Behaviour assessment system, for children teacher rating scales Classroom competence observation form, Observation of goal behaviours
Outcome sizes	
Blinding of Assessors	
Ppt Selection	Contacted directors of education
Main Investigation Question	Can training in FBA and positive behaviour support impact teacher self-efficacy, competence, and student behaviour?
Potential Limitations	Significantly higher change for target than general children - are staff generalising skills or only using for specific plans worked on in training? Knowledge of specific factors of training that lead to improvements is unknown
Reliability	

8. Revital, S.S. (2009). Dealing with school violence: The effect of school violence

prevention training on teachers' perceived self-efficacy in dealing with violent events. *Teaching and Teacher Education.* **25**(8), 1061-1066.

Participants	Anonymous questionnaire to 147 teachers
Intervention	School violence prevention training
Outcomes	Self-efficacy (teacher outcome, personal, and in school)
Design	Comparison of groups — Questionnaire
Comparisons	(41% of participants had participated in violence prevention training)
Results	Significant correlation between taking part in violence training and perceived outcome efficacy in dealing with violence ($p<0001$) (significantly higher in qualified teachers than student teachers, also teachers with more experience = higher self-efficacy for this) No change in personal self-efficacy or general self-efficacy
Outcome Measures	Developed own questionnaire incorporated-demographics, training information, and self-efficacy questions
Outcome sizes	
Blinding of Assessors	N/A
Ppt Selection	Questionnaire to schools in a locality that had received violence training
Main Investigation Question	School violence prevention training and how this correlates with teacher perceived self-efficacy about handling violent events
Potential Limitations	Self-report, unaccounted for variables (e.g. position in school)
Validity	No pre-measures

9. Liaw, E. (2009). Teacher efficacy of pre-service teachers in Taiwan: The influence of classroom teaching and group discussions. *Teaching and Teacher Education.* **25**(1), 176-180.

Participants	26 pre-service teachers
Intervention	Teacher preparation program – worked together as a group to review video recorded sessions of themselves

	and problem solve
Outcomes	Self-efficacy
Design	Questionnaire and interview. Post- intervention mixed methods
Comparisons	N/A
Results	After sessions: Self-efficacy to: - Motivate students (increased) - Deal with environmental factors (decreased) - Own ability to manage a classroom (increased) - Ability to select appropriate learning materials (increased) Interview common themes - Group discussions increased confidence in knowing what to do - Empathy between group members seen as very positive
Outcome Measures	Teacher efficacy scale (adapted) Interviews
Outcome sizes	Not significant (quantitative)
Blinding of Assessors	No
Ppt Selection	Enrolled on a course
Main Investigation Question	Do group discussions about classroom teaching impact self-efficacy of trainee teachers?
Potential Limitations	No pre-measures Questionnaire Compulsory part of course
Reliability	Results due to increasing experience or group discussions?

10. Hargreaves, D. J., Purves, R. M., Graham, F. (2007). Developing identities and attitudes in musicians and classroom music teachers. *British Journal of Educational Psychology.* **77**(3), 665-682.

Participants	29 trainee secondary music school teachers
Intervention	Trainee music students – (educations group) – received training during course and then opportunities to practise in post.
Outcomes	Changes in attitude, self-efficacy (music teaching) and professional belonging
Design	Longitudinal questionnaire study – during education and then first teaching posts
Comparisons	2 groups of 29 student teachers (one control (undergraduate music students), one experimental) as well as compared pre- and post- course
Results	ANOVA No significant interactions / effects in self-efficacy, professional belonging Significant difference in some attitudes – e.g education students (teacher course) significantly higher bias towards teaching rather than specialising in music also towards teachers emphasising social benefits of music. Non-education group thought teachers should emphasise intrinsic value of music (also this change occurred over time in education group)
Outcome Measures	Specially devised "Musical Careers Questionnaire" – incorporates questions around self-efficacy in music teaching, professional group identification, attitude
Outcome sizes	Power of ANOVAs all too small due to sample size
Blinding of Assessors	n/a
Ppt Selection	Contacted course providers for volunteers
Main Investigation Question	Does music teacher training course impact on self-efficacy for teaching, professional belonging and attitudes?
Potential Limitations	Music teaching – has a different research base due to specific nature and skill set therefore is this generalisable to other subjects? Education group – not all same education – just all on a

	training course for music teaching.
Reliability	Assumes musicians not in teacher training course are not getting opportunities to teach. Questionnaire (devised for purpose of this research)

11. Shechtman, Z. Levy, M., Leichtentritt, J. (2005). Impact of life skills training on teachers' perceived environment and self-efficacy. *Journal of Educational Research.* **98**(3), 144-154.

Participants	360 teachers from 360 5h and 6th grade classes (97 schools) in Israel. 214 used LST (life skills training)
Intervention	Teacher training in LST - psychoeducational model – focus on life skills – disseminated by teachers and school counsellors
Outcomes	Self-efficacy scores, work climate ratings, implementation variables (correlations and factor analyses)
Design	Questionnaire, - between groups, quantitative.
Comparisons	3 groups- no training (65), one year of training (84), two years of training (65)
Results	Self-efficacy – correlated significantly and positively with supervisor support and clarity of rules, negatively with work pressure. Significantly higher self-efficacy and positive work environment interpretations in 2 year training than 1 year and no training. Self-efficacy – correlated – professional lives and personal lives (implementation variables) Most implementation variables – related to work climate dimensions Mediating variables – SES and class size
Outcome Measures	Abridged work environment scale Teacher self-efficacy scale Teacher evaluation questionnaire (implementation variables)
Outcome sizes	N/A

Blinding of Assessors	N/A
Ppt Selection	Schools randomly selected from the 600 schools that provided LST training.
Main Investigation Question	As the LST is completed in a group and focuses on life skills will it lead to improved teacher perceptions of the work environment and higher levels of teacher self-efficacy?
Potential Limitations	- Retrospective - May be different quality training in different schools?
Reliability	- Self-report measures

12. Telljohann, S.K., Everett, S.A. Durgin, J. Price, J.H. (1996). Effects of an inservice workshop on the health teaching self-efficacy of elementary school teachers. Journal of School Health. 66(7), 261-265.

Participants	Elementary school teachers (262)
Intervention	30 hour health education training programme "Project Healthy Kids"
Outcomes	Health teaching self-efficacy, amount of time per week teaching health education, amount of effort in specific health ideas
Design	Matched groups – pre- and post-test
Comparisons	Experimental group (112), control (150) (same schools).
Results	Increase in experimental group compared to control for: self-efficacy ($p=0.03$) outcome expectations ($p=0.03$) outcome value ($p<0.001$) hours teaching health ed ($p=0.01$) subject time and effort ($p=0.08$)

	Experimental group – pre- and post- – improvement in: Efficacy expectation (p<0.001) Outcome expectation (p<0.001) Outcome value (p=0.001) Hours teaching health education (p=0.002) subject time and effort (p<0.001)
Outcome Measures	41 item survey created – addressed: self-efficacy constructs, outcome expectations, efficacy expectations, outcome value, health teaching time
Outcome sizes	
Blinding of Assessors	N/A
Ppt Selection	Volunteers for training and matched pairs
Main Investigation Question	Does in service training in health education impact efficacy and teaching practices of teachers?
Potential Limitations	Nonresponse bias? 52% return Convenience sample – one specific school district
Reliability	Own measure Self-report measures

13. Kaspereen, D. Relaxation Intervention for Stress Reduction Among Teachers and Staff. *International Journal of Stress Management.* **19**(3), 238-250.

Participants	54 teachers and staff from a high school
Intervention	Relaxation therapy – 30-35 min per week for 4 weeks (meditation, deep breathing and relaxing music)
Outcomes	Overall perceived stress, perceived work stress, life satisfaction
Design	RCT, pre- and post- comparisons
Comparisons	Intervention and waiting list control group
Results	Perceived stress increased for experimental group (p<0.001), control difference ((p=-0.36)

	Work stress decrease in the control group (p=-0.99) Experimental increased in life satisfaction (p<0.001)
Outcome Measures	Perceived Stress Scale Professional Life Stress Scale Satisfaction with Life Scale
Outcome sizes	
Blinding of Assessors	no
Ppt Selection	54 volunteers
Main Investigation Question	Relaxation therapy – lead to decreased overall stress, lowered work stress and increased overall life satisfaction
Potential Limitations	1 high school only Therapist and researcher same – response bias? Follow up – 1 week later only
Reliability	All aware of study and purpose of study Self-report surveys

Appendix 3: Pilot Evaluation Form

Pilot – Training in Managing Behaviour that Challenges

Date: 06.11.13

Your role _____

What went well
Would be better if

For me the **most** useful and valuable part of the training was…

For me the **least** useful and valuable part of the training was….

All feedback will be kept and recorded anonymously.

Please do not hesitate to contact me with any questions or comments about the training or the research project.

Thank you very much for your time

Heather Cooke

Trainee Educational Psychologist

Email: Heather.cooke@northyorks.gov.uk

Lpxhc2@nottingham.ac.uk

Phone: 01609797263

Appendix 4: Pilot Evaluation Data

What went well	**Number of staff**	**Would be better**	**Number of staff**
Lots of information/good content	10	Examples to support Information	1
Very useful	4	Behaviour plan example	10
Applicable behaviour plan/ discussion about plan/putting theory in to practice	5	Opportunities to discuss in groups/interactive sections	2
Learning about the theory behind behaviour	1	Didn't understand slide on functions	1
Reminding to record behaviour before using plan	1	Slower/ More time	2
Confident presentation/good pace/ friendly approach that welcomed questions	5	More description of the technical terms	1
Prompt questions on behaviour plan	1	More relevant to younger children	1
Incorporation of interactive activity at start	2		
Thinking about positive strategies	1		

Most useful	**Number of staff**	**Least useful**	**Number of staff**
Behaviour Plan	7	Feel more the domain of a teacher to teaching assistant	1
Thinking of behaviour as not only negative outbursts/ thinking of different ways of exhibiting challenging behaviour	2	Some of it was general and already known	1
Handouts	1	Behaviour definitions	1
Theory behind why children have challenging behaviour	1		
Being reminded of the little things	1		
Different ways of dealing with behaviour for different aged children	1		

Appendix 5: Behaviour Plan with Example Questions

Example questions – they do not all need to be asked – these are prompts if they are needed, ask the questions you think most appropriate for the child and the situation.

Name:

My behaviour plan

Picture/description of desired outcome in child's words

3. Aims:

Description of their desired behaviour/selves

- What will your behaviour look like?
- What will you be like?

Should be realistic and achievable
If I came in to class how would I know you were acting/behaving in … way?
Should be a specific and measurable visible change – describe what it will look like.

1. What does my behaviour look like now:

Who was there?
Where were you?
What did it look like?
What happened before/during after?
How did you feel before/during/after?
If I would have walked in to…. when….. what would I have seen?
What usually happens before? (If struggles if you went back to class now what would the teacher/peers etc. do that would make you feel like/act like that?)

2. What will my behaviour look like when my aims have been me?

What skills do you have?
What do you do to calm yourself down?
What do you do/what happens to stop it from getting worse
What do you want your behaviour to look like?
Who will be the first to notice?
What would be the first thing you/they would notice if it changed?
What is the first small steps
What changes do you need to make?
What can you think of that you might be able to do instead?
If you could have anything what would it look like?
What would your behaviour look like if that were the case?
What is different to now?

Name: Week commencing:

Weekly focus:

	9:00 – 9:30	9:30 – 10:30	Break	10:45 – 12:00	Lunch	1:00 – 2:15	Break	2:30 –3:15
Monday								
Tuesday								
Wednesday								
Thursday								
Friday								

Target: /40 Actual: /40

Reward:

Key

√ Focus achieved

O Focus not achieved

Appendix 6: Training Powerpoint for Schools

Challenging Behaviour in Schools

Heather Cooke and Matthew Davies

What is Challenging Behaviour?

Why?

Functions (Cooper, Heron & Howard, 2007)

- social attention,
- tangible reinforcement,
- self-stimulation,
- social negative reinforcement,
- automatic negative reinforcement

Attachment etc.

- Secure base

Chain of events

Breakwell, 1997 – the assault cycle

How do we manage challenging behaviour?

Rewards

- If these will be used they need to relate to the function AND be rewarding for the individual child

Replace the behaviour

- How can we expect a child to behave in a certain way if they do not have the skills required

Positive directions…

- There are millions of alternative behaviours to "do not (shout)" – which one should they choose? Make it clear by explicitly labelling the desired behaviour.

Consistency

- (and secure base)

Ethics

Learning behaviours

Modelling

Positive R+ of positive behaviours is more powerful and positive than punishing negative behaviours

What do we need to know?

(Rae & Daly, 2008)

What brings about the challenging behaviour and why?

The feelings they have when the behaviour (or pre-behaviour) occurs

What would the child like to happen?

What change is needed to get to their vision?

How they would feel if these changes were put in place

How will suggested changes assist in resolving their problem?

How do we find the function?

When is the behaviour happening?

Is there something that starts it in the environment? Reactions from others?

Data
- To understand but also to help monitor and set realistic targets that the child can see

Solution focused questioning
- What skills do I have?
- What do I hope to learn?
- What changes do I need to make?

Pupil led Target setting

Solution focused questions lead to targets and aims for the child

Don't forget the "what to do instead" question.

Regular feedback to the child on how they are doing increases their engagement (and ours)

Targets need to be realistic and achievable

Behaviour Plan

The plan incorporates:
- Weekly meeting with the child
- Finding the function, replacing the behaviour
- Solution focused questioning – give the pupil the lead to get to a more positive self
- Target setting
- Monitoring
- Regular feedback and rewards

What does my behaviour look like now?

Here we need to use questions that can help us find out when/where/ why (function):-
- Tell me about last time…
- Who was there?
- What happened before?
- What happened after?
- How did you feel?
- Does that usually/what do you/don't you like about …?
- Mirroring behaviours (make sure you get permission)

What will my behaviour look like when my aims have been met?

Use the function to guide this as well as solution focused questioning to find out the pupils strength and work to these.
- Questions:

Solution focused questioning
- What skills do I have?
- What do I hope to learn?
- What changes do I need to make?
- When I get mad it is usually because…. When I get mad I want to…
- When I get mad it helps me to calm down if…. After I have been mad I feel.. When I get mad my body feels….

Setting Aims

Description of their desired behaviour/selves

Realistic

Achievable

Relate to their chosen challenging behaviour

Specific and measurable!!

Target Sheet

Realistic achievable target

Child needs to experience success – this allows the child to see their success and take ownership of their own behaviour plan.

Think about red and green behaviours.

Aim for a gradual decrease over a number of weeks, do not expect too much too soon.

Reward – should be chosen by the child as far as possible – if the child does not find it rewarding it is not a reward!

Unless someone like you cares a whole awful lot,

nothing is going to get better.

It's not.

Dr Suess

Appendix 7: Pupil Behaviour Definitions

Pupil: S1a

Behaviour: Inappropriate reaction to not being chosen (hand up).

Definition: If another child is chosen to answer a question when S1a has put a hand up and he responds to this by verbally signalling his displeasure (more so than typical of peers).

Examples: S1a sighs loudly or complains verbally that he has not been chosen to answer the question.

Non-Examples: S1a looks disappointed when he is not asked to answer the question.

Behaviour: Interrupting Teacher.

Definition: The teacher / a staff member is speaking to the class / another child and S1a interrupts by answering when not appropriate or making comments without first putting his hand up / getting the staff members attention appropriately.

Examples: S1a shouts out an answer when a staff member asks the class and it is clear it is expected pupils put their hands up. S1a makes an uninvited comment when a staff member is addressing someone else / the whole class or when another pupil is addressing the class / answering a question.

Non-Examples: S1a answers a question after it is posed to him. S1a answers a question / speaks to a staff member / peer in a discussion.

Behaviour: Complaints about lack of help given from staff (Familiar work).

Definition: S1a complains that he is not being helped or supported by staff when completing a task that is within his ability to complete independently.

Examples: S1a says "I can't do it", "no one is helping me" or similar when completing a piece of work within his ability to complete without support.

Non-Examples: S1a appropriately asks for help with a piece of work that he finds challenging to complete after making some attempt at independently completing the task.

Behaviour: Refusal to complete work /t asks when supported.

Definition: S1a refuses to complete his work when he is being supported by a staff member.

Examples: S1a is receiving support from a staff member and refuses to complete work because they are not helping him "enough". Comments are similar to "You don't help me, I won't do it".

Non-Examples: S1a appropriately asks for help with a piece of work that he finds challenging. S1a stops work to wait for help / listen to support but continues willingly when he understands the tasks.

Pupil: S1b

Behaviour: Inappropriate crying.

Definition: S1b cries and complains loudly in a manner which is louder and longer than is expected from peers his age in response to similar events.

Examples: S1b stops doing what he is doing and cries loudly when a child brushes past his arm gently. S1b cries loudly if asked to complete a task he does not want to complete.

Non-Examples: S1b stops what he is doing and cries when he has been hurt.

Behaviour: Inappropriate noises during work time.

Definition: When S1b is supposed to be listening / quiet he makes loud vocalisations or noises.

Examples: S1b makes high pitched noises while the teacher is addressing the class / a group.

Non-Examples: S1b hiccups when he is supposed to be being quiet. S1b asks a question or contributes to a discussion. S1b makes high pitched sounds in his free time (e.g. lunch or break time)

Pupil: S1c

Behaviour: Refusal to comply with an instruction given by a staff member.

Definition: S1c refuses to comply with an instruction from a staff member after 5 seconds.

Examples: S1c refuses to move after being asked to complete something by a staff member, S1c refuses to stop what he is doing when asked to by a staff member, S1c completes different activities to the ones he is asked to complete by a staff member.

Non-Examples: S1c does not complete an instruction because he does not hear the instruction from a staff member. S1c begins completing an instruction but stops to ask for help as he cannot complete the request alone. S1c asks for clarification on a task within 10 seconds of being asked to complete the task.

Appendix 8: Example Behaviour Collection Sheet (S1a)

Date: _____

Name: _____

Target Behaviour: Please see definitions sheet for accurate description of behaviours

	Inappropriate reaction to not being chosen (hand up)	Interrupting Teacher	Complaints about lack of help given from staff (Familiar work)	Refusal to complete work/tasks
Example	*III*	*I*	*0*	*IIII*
Monday a.m (11.11.13)				
Monday p.m				
Tuesday a.m				
Tuesday p.m				
Wednesday a.m.				
Wednesday p.m				
Thursday a.m				
Thursday p.m				
Friday a.m				
Friday p.m				

Appendix 9: Ethics Committee Approval Letter

AS/hcf

Ref. 325

Monday, October 13, 2014

Dear Heather Cooke,

Ethics Committee Review

Thank you for submitting an account of your proposed research 'The impact of training in a pupil centred behaviour plan on staff self-efficacy, staff burnout, and pupil challenging behaviour'.

That research has now been reviewed, we are pleased to tell you it has met with the Committee's approval.

However:

Please note the following comments from our reviewers;

1. The letter to parents is too long and complex. It needs to be rewritten to concisely inform the parent what will happen to their child and when.

2. If the abbreviation SENCO is used in the revised letter then it needs to be defined.

3. "As part of this training I am hoping to carry out a doctoral thesis" is a bit weird. If this included in the revised letter, better would be to say for example that you hope to carry out this research project for your doctoral thesis.

4. "If you decide to participate key staff members…". It would be useful to make clear that staff members from the school of your child will be trained. I assume it is not so relevant to indicate when the training takes place but more important is for the parents/carers to know when exactly the intervention will take place.

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

Dr Alan Sunderland

Chair, Ethics Committee

Appendix 10: Information and Consent for School Staff (Research Question 1)

University of Nottingham

School of Psychology

Information Sheet

Training and Intervention for Behaviour That Challenges

Researcher: *Heather Cooke*

Contact Details*: lpxhc2@nottingham.ac.uk*

University Supervisor: *Neil Ryrie*

School of Psychology

University Park Campus

University of Nottingham

Nottingham

NG7 2RD

Placement Supervisor: XXX

Information for Schools

I am a trainee educational psychologist studying at Nottingham University and am currently on placement with the XXX Educational Psychology and Early Years Service. As part of the training, I am hoping to carry out a doctoral thesis researching the effectiveness of school staff training and individualized intervention for challenging behavior developed by the XXX Enhanced Mainstream School for Behavioural, Emotional and Social Difficulties (EMS, BESD).

This is an invitation to inform you of the research and to invite you to take part. The reason you have been approached with this opportunity is that XXX XXX, Specialist Teacher, BESD feels that schools within the XXX EMS area should be able to benefit from the strategies and training developed at the EMS with regard to behavior management. This will be an opportunity to receive training, and give feedback.

Before you decide if you wish to take part in this research project 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.

A wait-list control trial (participants will be split into groups and given the training at different times of the year according to group) will be used to evaluate the effectiveness of a training package developed by the XXXX Enhanced Mainstream School for Behavioural, Emotional and Social Difficulties, in conjunction with Heather Cooke, Trainee Educational Psychologist. Schools involved in the research will be randomly allocated to either the experimental or the wait-list control group. Whole school teaching staff training will be given in November/December 2013 (or after January 2014) as part of the requirements

of the research project. All staff members taking part in the research regardless of group allocation will be asked to complete questionnaires in November / December 2013 about the training and themselves and also in January 2014.

All questions are voluntary and questions can be left blank if staff do not feel they wish to answer. I will be available to support you and any staff members involved with any queries throughout the project via e-mail and telephone using the contact details below.

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.

If you have any questions or concerns please do not hesitate to contact me at any time. I can also be contacted after your participation at the above e-mail address.

Thank you for taking the time to read this information. I hope you will be interested in this training and research opportunity.

Heather Cooke

Trainee Educational Psychologist

Email: lpxhc2@nottingham.ac.uk

Phone: xxx

CONSENT FORM FOR SCHOOLS AND SCHOOL STAFF

Research considering Staff Training and Intervention for Challenging Behaviour

Researcher: Heather Cooke (Trainee Educational Psychologist, School of Psychology, University of Nottingham).

Please complete the whole of this sheet yourself. Please cross out as necessary.

- Have you read and understood the schools information sheet YES/NO
- Have you had the opportunity to ask questions and discuss the study YES/NO
- Have all the questions been answered satisfactorily YES/NO
- Have you received enough information about the study YES/NO
- Do you understand that you are free to withdraw from the study:

  at any time YES/NO

  without having to give a reason 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

Appendix 11: Information and Consent for School Staff (Research Question 1 and 2)

University of Nottingham

School of Psychology

Information Sheet

Training and Intervention for Behaviour That Challenges

Researcher: *Heather Cooke*

Contact Details*: lpxhc2@nottingham.ac.uk*

University Supervisor: *Neil Ryrie*

School of Psychology

University Park Campus

University of Nottingham

Nottingham

NG7 2RD

Placement Supervisor: xxx

YO12 6E

Information For Schools

I am a trainee educational psychologist studying at Nottingham University and am currently on placement with the XXX Educational Psychology Service. As part of the training I am hoping to carry out a doctoral thesis researching the effectiveness of school staff training and individualised intervention for challenging behaviour developed by the XXX Enhanced Mainstream Service for Behavioural, Emotional and Social Difficulties.

This is an invitation to inform you of the research and to invite you to take part. The reason you have been approached for this research is that XXX (Specialist Behavior Teacher) feels that schools within his area should be able to benefit from the strategies and training developed at the service. This will be an opportunity to receive training, have support in implementing the intervention with a number of students, and give feedback.

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.

A wait-list control trial (participants will be split into groups and given the training at different times of the year according to group) will be used to evaluate the effectiveness of a training package developed by the XXX Enhanced Mainstream School for Behavioural, Emotional and Social Difficulties, in conjunction with Heather Cooke, Trainee Educational Psychologist. Schools involved in the research will be randomly allocated to either the experimental or the wait-list control group. Whole school staff training will be given in November/December 2013 (or after January 2014) as part of the requirements of the research project. All staff members taking part in the research regardless of group allocation will be asked to complete questionnaires in November 2013

about the training and themselves and also in January 2014. All questions are voluntary and questions can be left blank if staff do not feel they wish to answer. I will be available to support you and any staff members involved with any queries throughout the project via e-mail and telephone using the contact details below.

You have also been chosen to take part in the part of research which looks directly at pupil behavioural outcomes of the intervention. This means in addition to attending the school training, staff members will be supported in implementing the behaviour plan with individual students. After training these staff will be supported in recording challenging behaviours on a daily basis and supporting a student using this intervention to decrease their challenging behaviour. Regular visits from myself will be arranged to support staff with this implementation and to collect regular data from staff.

I will be available to support you and any staff members involved with any queries throughout via e-mail and telephone using the contact details below.

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.

If you have any questions or concerns please do not hesitate to contact me at any time. I can also be contacted after your participation at the above e-mail address.

Thank you for taking the time to read this information

Heather Cooke

Trainee Educational Psychologist

Email: lpxhc2@nottingham.ac.uk

CONSENT FORM FOR SCHOOLS AND SCHOOL STAFF

Research considering Staff Training and Intervention for Challenging Behaviour

Researcher: Heather Cooke (Trainee Educational Psychologist, School of Psychology, University of Nottingham).

Please complete the whole of this sheet yourself. Please cross out as necessary

- Have you read and understood the schools information sheet YES/NO
- Have you had the opportunity to ask questions and discuss the study YES/NO
- Have all the questions been answered satisfactorily YES/NO
- Have you received enough information about the study YES/NO
- Do you understand that you are free to withdraw from the study:

  at any time YES/NO

  without having to give a reason YES/NO

- Do you agree to take part in both parts of the study (randomised control trial and individual student support) 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

Appendix 12: Information and Consent for Parents / Carers

University of Nottingham

School of Psychology

Information Sheet

Training and Intervention for Behaviour That Challenges

Researcher: *Heather Cooke*

Contact Details*: lpxhc2@nottingham.ac.uk*

University Supervisor: *Neil Ryrie*

School of Psychology

University Park Campus

University of Nottingham

Nottingham

NG7 2RD

Placement Supervisor: XXX

Information For Parents / Carers

I am a trainee educational psychologist studying at Nottingham University and am currently on placement with the XXX Educational Psychology and Early Years Service. As part of the training, I am hoping to carry out a doctoral thesis researching the effectiveness of school staff training and individualized intervention for challenging behavior developed by the XXX Enhanced Mainstream School for Behavioural, Emotional and Social Difficulties (EMS, BESD).

This is an invitation for your child to take part in a research study exploring the effects of a school-based intervention developed to support students to manage their own challenging behaviour. Your child has been chosen by the SENCO and their class teacher as a student who may benefit from this intervention, which will aim to support your child to practise their “best behaviors” and be rewarded for this regularly.

The reason you have been approached is to ensure you are able to give fully informed consent for your child. 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.

If you decide to participate key staff members will be trained in November / December 2013 in delivering the intervention. These staff will be asked to take daily data (number of times your child behaves well or in a challenging way) to inform the research. The intervention will then be put in place by the trained school staff who will receive regular support from myself (trainee educational psychologist).

The intervention will involve your child working 1:1 with a teaching assistant once a week where they will discuss their behaviours in school and set themselves achievable behaviour targets for the week. Your child will receive regular feedback from the class staff on how well he/she is meeting his/her behaviour targets. Your child and the teacher will take data on their behaviour and weekly rewards will be given dependent on this data. Following a 6 -10 week intervention period data will no longer be taken for purposes of the research. Staff and your child may wish to continue with the intervention, this will be supported by myself if necessary.

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.

If you would like to speak to me before consenting to take part in the research, or at any point during the research, I can be contacted through school or the details given below, and am happy to meet or call you at your convenience. If you have any questions or concerns please do not hesitate to ask. I can also be contacted after your participation at the e-mail address provided.

Thank you for taking the time to read this information.

Heather Cooke

Trainee Educational Psychologist

Email: XXX

lpxhc2@nottingham.ac.uk

CONSENT FORM FOR PARENTS / CARERS

Research Evaluating Intervention for Challenging Behaviour

Researcher: Heather Cooke (Trainee Educational Psychologist, School of Psychology, University of Nottingham).

Please complete the whole of this sheet yourself. Please cross out as necessary

- Have you read and understood the schools information sheet YES/NO
- Have you had the opportunity to ask questions and discuss the study YES/NO
- Have all the questions been answered satisfactorily YES/NO
- Have you received enough information about the study YES/NO
- Do you understand that you are free to withdraw from the study:

  at any time YES/NO

  without having to give a reason YES/NO

- Do you agree to take part in both parts of the study (randomised control trial and individual student support) 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

Appendix 13: Information and Consent for Students

Information Sheet and Consent for Participants (Students)

(*To be read with the teaching assistant/carer)*

Dear (insert student name),

I am writing to you to tell you about a research project you can take part in in school. This project aims to encourage pupils to think about their best behaviours and how to practise them in school. Sometimes pupils find school difficult and teachers want to be able to help them enjoy school more. Your teacher thought you might like some support to think about and work on your best behaviours so that you might start enjoying school more.

The work will involve:

- Working 1-1 with a teaching assistant for a short time every week away from the class
- Working together to talk about your behaviour and how you and the teaching assistant can help each other in school
- Having the teacher record your behaviour so you can talk about it together and
- Setting yourself targets (with your teaching assistant) for you to manage each week so you can choose a reward.

The way you behave in class will be recorded by the teacher and teaching assistant to be written about by the educational psychologist. All of this information would be kept very safe, and the written piece of work would not include your name. This means that anyone who reads the project would not know you were involved.

If you start these sessions and you or your family change your mind you can tell the teaching assistant. You would not be in any trouble for this and if you wanted to you could stop being part of the research. This would mean that no information about you would be included in the educational psychologist's work.

Your family know about the research project and if you want to you can discuss it with them before you decide if you want to take part. If you have any questions about this please ask the person who is reading this with you.

Consent

Have you read the information sheet? YES/NO

Do you understand what it means to be part of the project? YES/NO

Did you have a chance to ask questions? YES/NO

Were all of your questions answered? YES/NO

Do you understand that you can change your mind and stop being part of the project:

at any time YES/NO

without having to give a reason YES/NO

Would you like to be involved in the project? YES/NO

"This study has been explained to me and I would like to take part. I understand that I am free to leave at any time."

Signature of Pupil: Date:

Name (in block capitals)

I have ensured the study has been explained to the above pupil and he/she has agreed to take part.

Signature of researcher: Date:

Appendix 14: Example Questionnaire for School Staff

Research Study Questionnaire

Name: _____. School: _____

(please note, this will be removed immediately after the data collection and you will be assigned an anonymous code before any data analysis)

3. **Gender** (please tick the appropriate box)

Male ☐ Female ☐

4. Are you currently based in a class for most of the day? Yes / No
5. How many students in your class do you feel would benefit from a structured behaviour plan?_____
6. How many students in your class have a structured behaviour plan in place? _____
7. Are there students in your class who you feel may benefit from additional behaviour support? Yes / No
8. Have you implemented the behaviour plan from the training received in the previous half term (delivered by Heather Cooke) with any students in your class? Yes / No

If yes please state with how many students_____

If no please go to question 7.

6(b). If so do you feel this intervention has supported the student? Yes / No

9. Do you hope to use this intervention with students in your class? Yes / No

7b. If so with how many students?_____

10. Have you felt the information given in the training has supported you in your role?

8b. If yes how?

_____
_____
_____
_____

11. Do you have any comments you would like to make about the training?_____
    _____
    _____
    _____
    _____
    _____

School Staff Questionnaire 1

Please circle the number that best represents your view on whether you agree or disagree with each statement. You are asked to choose only one number from 1 to 6, with 1 being 'strongly disagree' with the statement and 6 being 'strongly agree' with the statement.

Item	Strongly disagree					Strongly agree
When a student does better than usual, many times it is because I exerted a little extra effort.	1	2	3	4	5	6
If a student becomes disruptive and noisy, I feel assured I know some techniques to redirect him/her quickly.	1	2	3	4	5	6
The hours I spend with a student have little influence compared to the influence of home environment.	1	2	3	4	5	6
I find it easy to make my expectations clear to students.	1	2	3	4	5	6
I know what routines are needed to keep activities running effectively	1	2	3	4	5	6
There are some students who won't behave no matter what I do.	1	2	3	4	5	6
I can communicate to students that I am serious about getting appropriate behaviour.	1	2	3	4	5	6
If one of my students couldn't do an assignment I would be able to accurately assess whether it was the correct level of difficulty.	1	2	3	4	5	6
I know what kinds of rewards to use to keep students involved.	1	2	3	4	5	6
If students aren't disciplined at home, then they aren't likely to accept it at school.	1	2	3	4	5	6
There are very few students that I don't know how to handle.	1	2	3	4	5	6
If a student doesn't feel like behaving there's not a lot teachers can do.	1	2	3	4	5	6
When a student is having trouble with an assignment, I am usually able to adjust to his/her level.	1	2	3	4	5	6
Student behaviour in the classroom is influenced more by peers than the teacher.	1	2	3	4	5	6
When a student gets a better grade than usual, it is probably because I found better ways of teaching that student.	1	2	3	4	5	6
I don't always know how to keep track of several activities at once.	1	2	3	4	5	6
When I really try I can get through to the most difficult students.	1	2	3	4	5	6

I am unsure how to respond to defiant students.	1	2	3	4	5	6
A teacher is very limited in what can be achieved because a student's home environment is a large influence on achievement.	1	2	3	4	5	6
I find some students to be impossible to discipline effectively.	1	2	3	4	5	6
When the grades of my students improve, it is usually because I found more effective teaching approaches.	1	2	3	4	5	6
Sometimes I am not sure what rules are appropriate for my students.	1	2	3	4	5	6
If a student masters a new concept quickly this might be because I knew the necessary steps in teaching the concept.	1	2	3	4	5	6
The amount that a student can learn is primarily related to family background.	1	2	3	4	5	6
I can keep a few problem student from ruining the entire class.	1	2	3	4	5	6
If parents would do more with their children at home, I could do more with them in the classroom.	1	2	3	4	5	6
If students stop working in class, I can usually find a way to get them back on track.	1	2	3	4	5	6
If a student did not remember information I gave in a previous lesson, I would know how to increase his/her retention in the next lesson.	1	2	3	4	5	6
Home and peer influences are mainly responsible for student behaviour.	1	2	3	4	5	6
Teachers have little effect on stopping misbehaviour when parents/carers don't cooperate.	1	2	3	4	5	6
The influences of a student's home experiences can be overcome by good teaching.	1	2	3	4	5	6
Even a teacher with good teaching abilities may not reach many students.	1	2	3	4	5	6
Compared to other influences on student behaviour, teacher's effects are very small.	1	2	3	4	5	6
I am confident in my ability to ensure that students will learn and behave well.	1	2	3	4	5	6
I have very effective behaviour management skills.	1	2	3	4	5	6

(Taken from Teacher Efficacy in Classroom Management and Discipline Scales (Emmer & Hickman, 1991))

School Staff Questionnaire 2

Please indicate how often, by circling the number that best represents your views in the past 30 workdays, you have felt each of the following feelings:

How often have you felt this way at work?							
	Never or almost never	Very infrequently	Quite infrequently	Sometimes	Quite often	Very frequently	Always or almost always
I feel tired	1	2	3	4	5	6	7
I feel like my batteries are dead	1	2	3	4	5	6	7
My thinking process is slow	1	2	3	4	5	6	7
I feel fed up	1	2	3	4	5	6	7
I have difficulty thinking about complex things	1	2	3	4	5	6	7
I feel I am not capable of investing emotionally in coworkers and students	1	2	3	4	5	6	7
I feel I am not thinking clearly	1	2	3	4	5	6	7
I feel physically drained	1	2	3	4	5	6	7
I have difficulty concentrating	1	2	3	4	5	6	7
I feel I am unable to be sensitive to the needs of coworkers and students	1	2	3	4	5	6	7
I feel I am not focused in my thinking	1	2	3	4	5	6	7
I feel burned out	1	2	3	4	5	6	7
I have no energy for going to work in the morning	1	2	3	4	5	6	7
I feel I am not capable of being sympathetic to co-workers and students	1	2	3	4	5	6	7

Taken from (Shirom-Melamed Burnout Meausre, 2005).

Thank you for your time,

Heather Cooke,

Please speak to/contact Heather Cooke if you have any concerns/questions about this questionnaire.

Email: XXX

Phone: XXX

Appendix 15: Intervention Fidelity Treatment Check

Fidelity Check

Date:

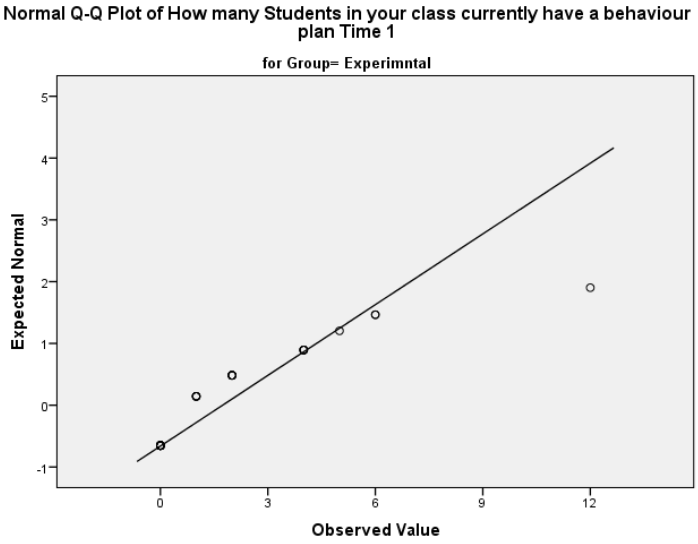
School:

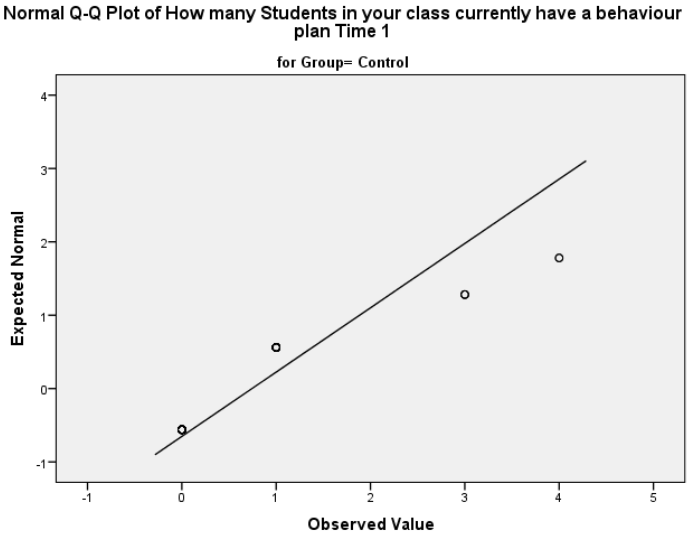
Pupil:

	Present	Comments
Completed data Sheet with clear target and reward system (after first week).		
Positive judgement free discussion		
Discussion of successes from last week		
Discussion about what the behaviour is – function based		
Solution focused discussion – building and focusing on the child's strengths		
Clear description of desired behaviour – based on function and strengths from previous discussion.		
Obvious link between discussion and goal		
Goals and target setting led by the child wherever possible (use of child's own words)		
Observable and measurable target and goals. Clearly described.		
Achievable target		

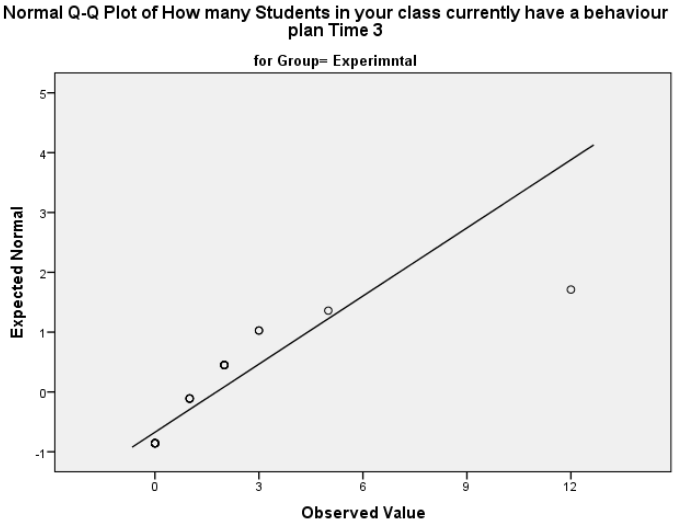
Appendix 16: Histograms and Q-Q plot SPSS Output for Research Question 1

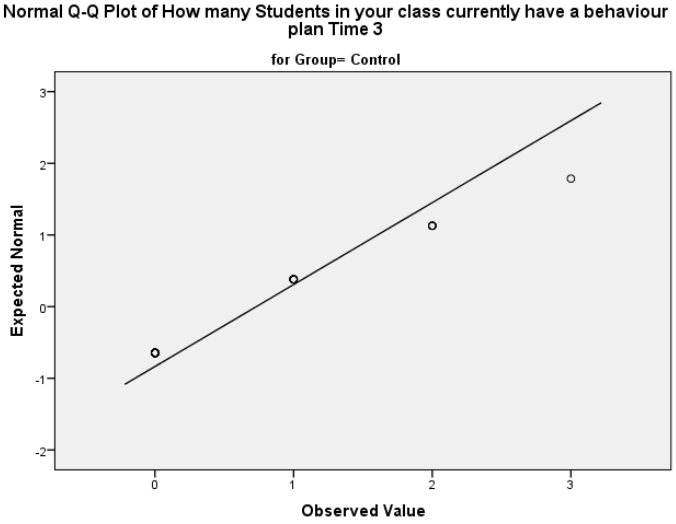
Time 1





Time 3


Normal Q-Q Plot of How many Students in your class currently have a behaviour plan Time 3
for Group= Experimntal
Expected Normal
Observed Value
5
4
3
2
1
0
-1
0
3
6
9
12


Normal Q-Q Plot of How many Students in your class currently have a behaviour plan Time 3
for Group= Control
Expected Normal
Observed Value
3
2
1
0
-1
-2
0
1
2
3

Appendix 17: Tables Showing Results of Tests of Normality (Shapiro-Wilks) for Self-Efficacy and Burnout Scores (Research Question 1b & 1c)

A table to show results for the tests of normality (Shapiro-Wilks) for actual Scores.

Dependent Variable	**N**	**Shapiro Wilks Result (Experimental)**	**N**	**Shapiro Wilks Result (Control)**
Time 1				
Self – Efficacy Overall	**28**	**D(28)=0.96, p=0.33**	**27**	**D(27)=0.96, p=0.30**
Self-Efficacy Personal	45	D(45)=0.97, p=0.35	54	D(54)=0.96, p=0.07
Self-Efficacy General	28	D(28)=0.98, p=0.83	31	D(31)=0.97, p=0.64
Self-efficacy External	44	D(44)=0.96, p=0.18	55	D(55)=0.98, p=0.42
Burnout Overall	**44**	**D(44)=0.86, p<0.01**	**56**	**D(56)=0.96, p=0.06**
Burnout Physical	47	D(47)=0.97, p=0.24	56	D(56)=0.96, p=0.04
Burnout Emotional	47	D(47)=0.72, p<0.01	56	D(56)=0.77, p<0.01
Burnout Cognitive	46	D(46)=0.91, p<0.01	57	D(57)=0.96, p=0.09
		Time 2		
Self – Efficacy Overall	24	D(24)=0.88, p=0.01		
Self-Efficacy Personal	27	D(27)=0.95, p=0.30		
Self-Efficacy General	24	D(24)=0.95, p=0.32		
Self-efficacy External	27	D(27)=0.97, p=0.61		
Burnout Overall	27	D(27)=0.98, p=0.82		
Burnout Physical	27	D(27)=0.98, p=0.82		
Burnout Emotional	27	D(27)=0.97, p=0.47		
Burnout Cognitive	27	D(27)=0.84, p<0.01		
		Time 3		
Self – Efficacy Overall	33	D(33)=0.98, p=0.71	36	D(36)=0.97, p=0.40
Self-Efficacy Personal	35	D(35)=0.98, p=0.70	37	D(37)=0.98, p=0.72
Self-Efficacy General	33	D(33)=0.96, p=0.19	36	D(36)=0.98. p=0.86
Self-efficacy External	35	D(35)=0.98, p=0.85	37	D(37)=0.96, p=0.19

Burnout Overall	35	D(35)=0.99, p=0.96	34	D(34)=0.98, p=0.68
Burnout Physical	35	D(35)=0.97, p=0.57	34	D(34)=0.97, p=0.41
Burnout Emotional	35	D(35)=0.88, p<0.01	34	D(34)=0.90, p<0.01
Burnout Cognitive	35	D(35)=0.97, p=0.44	34	D(34)=0.96, p=0.32

A table to show results for the tests of normality (Shapiro-Wilks) including estimated scores.

Dependent Variable	**N**	**Shapiro Wilks Result (Experimental)**	**N**	**Shapiro Wilks Result (Control)**
Time 1				
Self – Efficacy Overall	48	D(48)=0.98, p=0.63	59	D(59)=0.98, p=0.70
Self-Efficacy Personal	48	D(48)=0.97, p=0.36	59	D(59)=0.96, p=0.6
Self-Efficacy General	48	D(48)0.97, p=0.36	59	D(59)=0.54, p<0.05
Self-efficacy External	48	D(48)=0.97, p=0.18	59	D(59)=0.98, p=0.33
		Time 2		
Self – Efficacy Overall	27	D(27)=0.91, p<0.05		
Self-Efficacy Personal	27	D(27)=0.95, p=0.21		
Self-Efficacy General	27	D(27)=0.93, p=0.05		
Self-efficacy External	27	D(27)=0.98, p=0.84		
Self – Efficacy Overall	35	D(35)=0.98, p=0.71	37	D(37)=0.97, p=0.53
Self-Efficacy Personal	35	D(35)=0.98, p=0.70	37	D(37)=0.98, p=0.72
Self-Efficacy General	35	D(35)=0.96, p=0.22	37	D(37)=0.98, p=0.82
Self-efficacy External	35	D(35)=0.98, p=0.85	37	D(37)=0.96, p=0.19

Red numbers show significant results (p<0.05) and therefore do not meet the assumptions of normal distribution.

Appendix 18: Tables Showing Additional Data Collected from Participants.

Comments from staff (when visiting schools)
Teacher met me when I was in school (2 weeks after delivering training)– told me she had used information form the training to talk to a student about their behaviour and this had helped her understanding of the student as well as decrease the behaviour.
Teacher of S1c – Reports that she feels number of incidents is decreasing but not the amount of time spent refusing (Discussion after fidelity check 3)
Teacher of S1a and S1b – did not feel behaviours were altering but felt that her understanding of and relationship with S1b was improving as a result of the intervention (Discussion after fidelity check 2)
Teacher of S2a reports that she felt although incidents were very rare that she understands S1a's behaviour better (and believes S2a also does). She also felt that although data was non-conclusive for S2a (due to low frequency) that the incidents were shorter and S2a was coping with them better after they had occurred leading to less repercussions if the behaviour did occur. (Discussion after final fidelity check)
Teacher of S2b reports that she enjoyed the discussion with S2b and felt it had improved his behaviour in class very quickly.

Question	Experimental Group	Number of staff	Control Group	Number of staff
TIME 1	Phone call / email– are questionnaires needed to be completed by TA's? Questioned relevance to them	2	Arrived to collect questionnaires, no TAs completed as HT thought behaviour questions were not relevant to their role	1
TIME 3				
Have you felt the information given in the training has supported you in your role?	Yes	21		
	No	5		
If so how?	Increased understanding of	2		

	some behaviours / how to deal with challenging behaviours			
	Helped to make clear individual provision maps	1		
	Increased awareness about thinking about reasons for challenging behaviour	2		
	Helpful additional information	2		
	Secure strategies and process to follow (increased confidence in this)	5		
	Interested to hear different points of view	3		
	Emphasis on appropriate discussion with student is good – if time allows!	1		
	Taught me to break an issue down into small chunks to help manage it better	1		
	Increased number of strategies	2		
	It has made me realise that the children can't always help their behaviour and we can't expect them to change overnight	1		
	Open discussion regarding relevant curriculum and teaching methods for some children	1		
	Confidence to try new ways	1		
	New transferable ideas	1		
	If I need to implement a plan I know how	1		
	As Head I am encouraging staff to use this approach – also have said I am willing to conduct the process with individual children if necessary	1		
	Clear, practical support	1		

	Revised and refreshed input very helpful	1		
	Good basis for shared discussion with mutual support from each other	1		
	Clarity in approach and staff roles when implementing	1		
	Useful to support speaking to children about their targets and how to conduct conversations.	1		
	Being more tolerant by looking for signs of triggers for behaviour, being able to pre-empt and deal with inappropriate behaviours more easily.	1		
Do you have any additional comments you would like to make?	I believe training for behaviour needs to be tailored to the individual child	1		
	Too research based and not appropriate for classroom	1		
	Not different to what I already do / similar to current strategies	3		
	Not implemented as have a plan I already implement	1		
	Supported school ethos	1		
	Liked support paperwork	1		
	Do not have time to compete questionnaires multiple times	2		
	Do not have time to complete questionnaires not relevant to teaching role	1		