An Evaluation of the Effectiveness of 'Comic Strip Conversations' for Addressing the Target Social Behaviours of Primary-Aged Pupils on the Autistic Spectrum

Joanne Page

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

May 2016

Abstract

Educational provision for children with autism is increasingly being made within mainstream settings and a range of intervention strategies to cater for the diverse needs of this heterogeneous population are needed (Ali & Frederickson, 2006). This research presents an evaluation of 'Comic Strip Conversations' (CSCs) (Gray, 1994b) for addressing the target social behaviours of five primary-aged pupils with autism in mainstream schools. CSCs are a story-based intervention which use visual systems designed to support understanding of situations and encourage more appropriate social behaviours in individuals with autism. A systematic review of existing research into the effectiveness of CSCs highlights the limited evidence base that currently exists.

A series of multiple-baseline across behaviours single-case experimental designs (SCEDs) were implemented for four participants, in which two specific behaviours were targeted through a CSC intervention. An A-B SCED was implemented for a fifth participant, targeting a single behaviour. Repeated measures were taken through structured observations to assess the frequency of target behaviours. These measures were triangulated with pre- and post-measures of staff perceptions of the target behaviours and intervention effectiveness. This research additionally explored the relative impacts on behaviours of creating single versus multiple CSCs.

The repeated measures data was analysed using a combination of visual analysis and effect size analysis (Tau-U). The outcomes of this indicated mixed results, with the intervention appearing to be moderately to highly effective in addressing at least one target behaviour for three of the five participants. Outcomes in terms of changes in staff perceptions of target behaviours and ratings of intervention effectiveness were similarly mixed and did not consistently triangulate with the repeated measures data. The behaviour targeted through multiple CSCs demonstrated greater improvement than the behaviour targeted through a single CSC in three out of four participants,

however the difference was negligible in one case. Therefore, no clear association between intervention frequency and outcome could be concluded.

The results are considered in view of the limitations of the research, taking into account the research design, characteristics of the data obtained, and threats to internal validity. Implications for practice are outlined and suggestions are made for future research. The research concludes with some support for CSCs as a promising intervention which may aid the development of socially appropriate behaviours for some pupils with autism.

Acknowledgements

The author would like to thank:

The participants in this research and the staff who worked so hard to implement the intervention.

Her fellow trainees at the University of Nottingham and the course tutors for their support and encouragement. I owe particular thanks to my academic tutor, Dr Sarah Atkinson, for her support and guidance throughout the research process.

The staff from the Educational Psychology Service in which I was on placement for their encouragement and well wishes throughout, and the Autism Outreach Team staff who provided support for, and interest in, the research.

Her family and friends for their encouragement and understanding and, lastly, my partner Jon for his unwavering patience and good humour.

Table of Contents

List of Tables

List of Figures

List of Graphs

List of Acronyms

ADHD	Attention Deficit Hyperactivity Disorder
AET	Autism Education Trust
AOT	Autism Outreach Team
APA	American Psych
ASC	Autistic Spectrum Condition
ASD	Autistic Spectrum Disorder
BPS	British Psychological Society
CSC	Comic Strip Conversation
CSCS	Comic Strip Conversation – Single
CSCW	Comic Strip Conversation – Weekly
DfE	Department for Education
DSM-V	Diagnostic and Statistical Manual – 5th Edition
EP	Educational Psychologist
ERIC	Education Resources Information Centre
IOA	Inter-Observer Agreement
ITSA	Interrupted Time Series Analysis
PDD-NOS	Pervasive Developmental Disorder – Not Otherwise Specified
PEM	Percentage of Data Exceeding the Median
PND	Percentage of Non-Overlapping Data
RCT	Randomised Controlled Trial
SCED	Single Case Experimental Design
TB	Target Behaviour
WoE	Weight of Evidence

Chapter 1 -   Introduction

1.1   Background to this Research

The aim of the present research was to evaluate the effectiveness of 'Comic Strip Conversations' (CSCs), an intervention designed to support social skill development in individuals with autism and other social communication and interaction difficulties (Gray, 1994b). This research investigated the impact of the intervention on specific target behaviours and the relative impacts of the intervention in terms of frequency of use. As an evaluative study, the present research allies itself with the evidence-based practice agenda within the profession of educational psychology (Frederickson, 2002) and the recognised need for more rigorous evaluation of interventions at both group and individual level (Frederickson & Miller, 2008).

School can be a challenging environment for children with autism (Jones, 2002) and the difficulties experienced by these children have important implications for behavioural, language and social outcomes (Hutchins & Prelock, 2008). With provision for children with autism increasingly being made within mainstream settings (Dockrell, Ricketts, Palikara, Charman & Lindsay, 2012), staff in these settings need access to, and understanding of, effective interventions which foster the development of social integration, behaviour and well-being (Ali & Frederickson, 2006; Dockrell et al, 2012).

The research was conducted by a trainee educational psychologist as part of her professional training at the University of Nottingham. The researcher's interest in supporting pupils on the autistic spectrum had initially developed during previous experience of working with this population in both mainstream and special school settings. The researcher's interest in and knowledge of this area has subsequently been built upon through university and field-based experiences during her educational psychology training.

The research was conducted within the local authority in which the researcher was on placement. Within this authority, it was known that CSCs were sometimes recommended to schools as an intervention for pupils with social

communication and interaction difficulties by educational psychologists (EPs) and Autism Outreach staff, often alongside more well-known interventions such as Social Stories (Gray, 1994a). This led the researcher to question the evidence base of this intervention, which ultimately led to the development of the present research. It was discovered that only a limited amount of research into the effectiveness of CSCs exists, with very little of this being experimental in design. A unique contribution was made to the evidence base through this research being the first experimental study to evaluate the intervention with primary-aged pupils within mainstream schools in a UK context, and the first to explicitly explore potential differential effects of intervention frequency.

1.2   A Note on Terminology

There is, amongst some circles, currently a debate about whether the term 'autistic spectrum disorder' (ASD) or 'autistic spectrum condition' (ASC) is most appropriate to use. Some have argued in favour of 'ASC' as it is felt to be less stigmatising and better reflects the fact that individuals on the spectrum can have areas of cognitive strength, not just disabilities (Baron-Cohen et al, 2009). Jordan's (2007) discussion of the issue balances the need to have respect for the wishes of the more able individuals within the spectrum, while not depriving others of important resources that may depend on the acknowledgment of a 'disorder'.

The Autism Education Trust (AET) (2008) suggest that, in an ideal situation, the term ASC should apply to those individuals whose development and functioning is not significantly compromised (as long as their differences are still responded to appropriately) whereas the term ASD would then appropriately denote a 'true' disorder where development and functioning are compromised in even the most accommodating environments. However, the AET (2008) also argue that changing labels will achieve little if attitudes and resource allocation remain unchanged.

After considering the current debates, the author decided to use the term ASC rather than ASD throughout this paper to reflect the changing perceptions within society. In addition, the term ASC is now used by the key service

involved in conducting diagnostic assessments in the local authority in which the researcher was on placement. It was deemed important that the terminology used in this paper reflected the values of the context in which the research was conducted. The term 'autism' is also used interchangeably.

1.3   Structure of the Thesis

- Chapter 2 provides an overview of the key theoretical ideas and pertinent research which informed the research questions and the rationale for these.

- Chapter 3 explores various methodological issues and presents an account of the design and methods used in this research and the rationale for these.

- Chapter 4 provides discussion of the different methods available for analysing SCED data and details of the approach used in this research. The findings obtained from the five cases are presented, along with details of inter-observer agreement and intervention fidelity.

- Chapter 5 provides a discussion of the findings and possible links to the literature explored in Chapter 2. Limitations of the research and implications of the findings are discussed. Possible directions for future research and personal reflections are offered.

- Chapter 6 provides an overall summary, highlights the unique contribution made and presents final conclusions.

Chapter 2 - Literature Review

2.1 Introduction to the Literature Review

The literature review commences with an overview of current understandings of ASC, followed by consideration of some of the issues surrounding the educational provision provided for children and young people with ASC. A systematic review of the literature evaluating the effectiveness of CSCs is then presented.

2.2 The Autistic Spectrum

The label 'autistic' was first introduced in 1911 by the psychiatrist Eugen Bleuler and it originally referred to a particular disturbance in schizophrenia, namely the withdrawal of an individual from the outside world into the self (Frith, 2003). The first accounts of autism as a 'condition' in itself were published by the child psychiatrist Leo Kanner in 1943 (in Frith, 2003). In the detailed case study accounts of the 11 children he studied, Kanner described their characteristic features as being:

• autistic aloneness

• desire for sameness

• islets of ability

Soon after, the Austrian paediatrician Hans Asperger (1944, in Frith, 2003) published a paper entitled 'Autistic Psychopathies in Children'. Citing similarities as well as differences to the characteristics described by Kanner, Asperger described the peculiarities of communication, difficulties in social adaptation, movement sterotypies and unusual patterns of intellectual achievements (Frith, 2003).

These detailed case study accounts of children marked the first theoretical attempts at explaining the condition and, although being based on limited samples, marked an important first step in developing a diagnostic criteria for autism. The features identified in the 1940s are still considered to hold true to

this day although variation in the details has since developed, along with the addition of further key characteristics (Frith, 2003).

2.2.1 A Triad of Impairments

Wing & Gould (1979) carried out a landmark study in a London borough which aimed to find out how often each of the 'symptoms' previously identified by Kanner and Asperger were present in a population of children identified as 'handicapped', independently of any prior diagnosis. A group fitting Kanner's criteria were identified, as were some children fitting Asperger's criteria. There remained a number of children who did not fit either criteria but who had all kinds of combinations of features of these 'syndromes' which appeared to occur in a wide range of manifestations. What appeared to hold all these groups together was a 'triad of impairments' in social interaction, communication and imagination (see Figure 1).

Wing & Gould (1979) found that there was a strong tendency for these impairments to cluster together and it was difficult to draw neat boundaries between the named 'syndromes' and those with the triad of impairments who did not fit into either category. It was concluded that the concept of a spectrum fitted the findings better than the previously adopted categorical approach. The idea of a spectrum encapsulates the wide variation between individuals within each element of the triad. The category of Asperger syndrome within the spectrum is now well established (Frith, 2003) and applies to individuals with average or above-average intellect and good spoken language (Jones, 2002). A further subgroup is 'Pervasive Developmental Disorder - Not Otherwise Specified' (PDD-NOS), a term more commonly used in America which is often used when children do not meet all the criteria for ASC (Jones, 2002).

Social Interaction

➢ apparent aloofness and indifference to other people
➢ being interested in others in order to have their needs met
➢ acceptance and enjoyment of social contact, but inability to initiate and maintain it
➢ inappropriate or odd approaches to others
➢ stilted and overly formal interaction, even in the most able individuals
➢ limited understanding of unspoken social rules

Lack of Imagination & Rigidity of Thought

➢ adherence to repetitive play activities
➢ using play materials in an unusual manner (e.g. becoming preoccupied with an irrelevant detail)
➢ engagement in stereotypical behaviour (e.g. spinning, rocking, making noises)
➢ narrow range of interests
➢ difficulty with planning and organisation
➢ difficulty with change in routine
➢ lack of appreciation of consequences of their own actions and problems with the concept of cause and effect

Language and Communication

➢ lack of communication or difficulties in using language to communicate effectively
➢ difficulty with social aspects of language (e.g. turn-taking and timing)
➢ inappropriate social communication (e.g. shouting out or interrupting)
➢ inability to 'read' or understand the significance of facial expression, vocal intonation or body language
➢ inability to empathise with others
➢ literal understanding of language

Figure 1: The 'Triad of Impairments'. Adapted from Hannah (2001) and Ali & Frederickson (2006)

2.2.2 Diagnostic Criteria

The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V) (American Psychiatric Association, APA, 2013) is the current common world system of classification and diagnosis. Its diagnostic criteria for 'Autistic Spectrum Disorder' (ASD), as outlined below, continues to be underpinned by the triad of impairments:

1. Persistent deficits in social communication and social interaction across multiple contexts

2. Restricted, repetitive patterns of behaviour, interests, or activities

3. Symptoms must be present in the early developmental period (but may not become fully manifest until social demands exceed limited capacities, or may be masked by learned strategies in later life)

4. Symptoms cause clinically significant impairment in social, occupational, or other important areas of current functioning

5. These disturbances are not better explained by intellectual disability or global developmental delay.

In the DSM-V, the categories of Asperger syndrome and PDD-NOS are now included within the umbrella term 'ASD'.

The triad of impairments and the DSM-V criteria clearly illustrate the characteristic features of autism and demonstrate the difficulties and needs of the participants in the present research.

2.2.3 Prevalence

In recent years, there has been an increase in the number of children being diagnosed with ASC. Various reasons for this have been proposed including broader classification systems, increased awareness amongst practitioners, better identification and more sensitive assessment instruments (Wing & Potter, 2002).

Prevalence figures amongst the school-age population have been estimated to be approximately 1 in 100 (Baird et al, 2006; All Party Parliamentary Group on Autism, 2012). A potentially higher prevalence rate - 157 per 10,000 - has been estimated in one research report which additionally considered the ratio of known:unknown (considered to be about 3:2) cases (Baron-Cohen et al, 2009). However, identifying an accurate prevalence rate can be difficult as there is no register or exact count kept and any information is based on epidemiological surveys (National Autistic Society, no date).

ASC is three to four times more common in males than in females (APA, no date) and males particularly predominate at higher levels of ability (Frith, 2003). This observation remains unexplained, but the possibility that females are less likely to be diagnosed due to the triad of impairments, when combined with good language and high ability, being better masked in females (due to enhanced compensatory learning) has been considered (Frith, 2003).

2.2.4 Causes of Autism

It is now well-established that autism is a neurodevelopmental disorder with a biological basis and that genetic factors are strongly implicated (Medical Research Council, 2001). It has also been proposed that environmental risk factors may interact with gene susceptibility to trigger ASC or affect its severity, though neither the risk factors nor specific genes have yet been identified and the mechanisms that underlie the condition are still not well understood (Volkmar, 2011).

2.3 Psychological Theories of Autism

A number of psychological theories have attempted to explain the behaviours commonly seen in ASC and how individuals perceive, process and understand the world (Jones, 2002).

2.3.1 Theory of Mind

Of particular relevance to the focus of the present research is the 'theory of mind' hypothesis (Baron-Cohen, Leslie & Frith, 1985). This is a key cognitive theory of autism which suggests that many of the characteristics of autism stem from an impairment in the ability to 'mentalise'. 'Mentalising' is defined as the ability to attribute thoughts, beliefs and motives to others and understand that others have perspectives which may be different from our own. Such 'mind-reading' generally develops naturally in typically developing children without needing to be explicitly taught.

This theoretical concept has been tested through experimental investigations in laboratory settings using 'false belief' tasks, such as the 'Sally Anne' task (see Appendix 1). Baron-Cohen et al (1985) compared the performance of children with autism, children with Down syndrome and typically developing children on the Sally-Anne task. They found that 80% of the autistic children answered incorrectly, whereas 85% of the children with Down syndrome and 86% of the typically developing children answered correctly. Such findings have since been replicated in a number of studies, the majority of which have reported fail rates of above 70% for children with autism. This figure indicates that some children with autism *can* pass theory of mind tasks and this has been linked to higher verbal age (Happé, 1995). Difficulties in mentalising have been identified in everyday life as well as laboratory situations (Frith, Happé & Siddons, 1994). Further, research has suggested that a theory of mind impairment is specific to the condition to autism (Baron-Cohen, 1989; Frith et al, 1994; Leslie & Frith, 1988).

Despite the common focus on particular aspects of theory of mind (such as false beliefs), it is said to be more appropriately conceptualised as a broad and multifaceted construct (Astington & Baird, 2005). There is thought to be a wide array of constructs that may be subsumed under, or closely connected to, theory of mind, including affect recognition, visual perspective taking, empathy and the understanding of mental state terms (Hutchins, Bonazinga, Prelock & Taylor, 2008).

Experimental research (albeit on a small scale) has produced findings that indicate that using thought bubbles as a concrete, observable representation of intangible mental states can enhance autistic children's understanding of the thoughts of others. Parsons & Mitchell (1999) compared typically developing children, children with non-specific learning disabilities and autistic children's understanding of mental representations on a standard false belief task and another false belief task utilising the pictorial convention of thought bubbles. They found that children with autism (both higher and lower functioning in terms of verbal age) seemed to understand thought bubbles as representational devices and the inclusion of them in a task improved their performance. Similar findings have been reported by Kerr & Durkin (2004) and Wellman et al (2002), although the latter study found that although children with autism could understand and make use of the strategy, training was usually only associated with modest generalisation to transfer theory of mind tasks, even when these were similar to the training tasks. However, no transfer tasks involved assessing children's functioning in naturalistic situations.

It has been suggested that perhaps using pictorial cues such as thought bubbles enhances performance by reducing the peripheral demands of a task, for example, by lowering working memory demands (Kerr & Durkin, 2004). Some have argued that autistic children's difficulties with theory of mind tasks could be explained by working memory deficiencies (e.g. Gordon & Olson, 1998; Hughes, 1998).

2.3.2   Executive Dysfunction

There are some characteristics of ASC that are not readily explained by the delay or absence of theory of mind abilities, for example, restricted and repetitive behaviours and interests, inflexibility and difficulties with planning. It has been proposed that deficits in executive functions, which include abilities related to planning, attentional control and impulse control, can account for these characteristics (Ozonoff, 1997). Whilst executive dysfunctions are very frequently present in children with ASC, they are not unique to them as they are also found in other populations, such as those with Attention Deficit

Hyperactivity Disorder (ADHD) or those who have experienced traumatic brain injury (Pennington & Ozonoff, 1996).

2.3.3 Weak Central Coherence

Neither executive dysfunction nor theory of mind deficits can account for the relative strengths or special abilities that are demonstrated by some individuals with ASC. Frith (1989) suggested that such factors could be explained by weak central coherence. Central coherence refers to an inclination to assimilate information across stimuli to form coherent 'wholes' and to generalise across contexts in order to make sense of disparate inputs (Frederickson & Cline, 2009). Frith (1989) proposed that individuals with autism have a weak capacity for central coherence, which can sometimes have advantages. For example, children with ASC tend to show relatively superior performance on block design tests compared to typically developing children as they are better able to dissociate the parts from the whole.

There are implications of weak central coherence for education, particularly when it comes to incidental learning and generalisation. Therefore children with ASC should not be expected to apply previous learning in new situations without training or specific prompting (Frederickson & Cline, 2009).

2.4 From Theory to Intervention

The theoretical explanations of ASC can aid understanding of the difficulties that these individuals can have. The theory of mind account of autism in particular helps to explain many of the characteristic social and communication difficulties and differences and allows for a deeper understanding of these. As a result of increased understanding, it is possible to make allowances and provide appropriate support for individuals (Frith, 2003). Specific interventions shall be discussed later in this chapter.

2.4.1 Educational Provision for Children with ASC

School can be a challenging environment for children with ASC due to the level of social demands and potential for sensory overload (Jones, 2002). Understanding the actions of school staff and other children and participating in class can be challenging for these children. In particular, the ASC-specific impairments in theory of mind can prove a challenge to educational settings and they have important implications for behavioural, language and social outcomes (Hutchins & Prelock, 2008).

Educational provision for children with autism is increasingly being made within mainstream settings or, in some cases, specialist provision within mainstream schools, such as ASC resource bases (Dockrell et al, 2012). Despite increasing inclusion in mainstream schools, physical integration does not necessarily foster social integration, and the latter should be seen as a goal in itself and be actively targeted for intervention (Rogers, 2000).

A report by the Department for Education (DfE, 2012) stated that, despite a wealth of research examining the cognitive and behavioural profiles of pupils with autism, there is relatively little known about the needs of these pupils in mainstream classrooms and the ways in which these needs are met. The findings of the three-year prospective study identified that resources need to be targeted according to the particular social communication needs of individuals. Additionally, the report argued that schools need to be aware of the potential wider impacts of social communication difficulties on wellbeing, behaviour and peer relationships and should consider explicitly addressing these.

The teaching techniques that generally work well with other children are often found to fail to work for children with ASC (Jones, 2002) so there is a need to develop a range of intervention strategies to cater for the diverse needs of such a heterogeneous population of children (Ali & Frederickson, 2006). In addition, individuals will change over time so their needs will need to be reviewed regularly and responded to quickly (AET, 2008).

2.4.2 Interventions

The social difficulties in autism are probably the most defining feature of ASC (Jones, 2002). With social communication deficits being an early indicator and a key factor in long-term outcomes, this domain represents an important intervention target (Anagnostou et al, 2015). From the age of around five, autistic children usually demonstrate an improvement in social skills and general adaptation which generally continues throughout the rest of their development. However, a lack of 'mentalising' ability means that learning to behave appropriately with and towards others can be very slow (Frith, 2003). Improved social functioning has long been considered to be one of the most important intervention outcomes and children with ASC, whilst demonstrating primary difficulties in social interactions, can be responsive to a variety of interventions aimed at increasing their skills in this area (Rogers, 2000).

2.4.3 Social Skills

A number of approaches have been developed which target the social development of pupils with ASC. One such (relatively broad) approach has been social skills programmes. These can take a variety of forms but are generally carried out with small groups (which often include peers without ASC). For higher-functioning children who have ASC, cognitive-behavioural strategies are commonly used to help train social problem-solving, emotional understanding and social interaction skills (Frederickson & Cline, 2009).

Social skill programmes which specifically target 'theory of mind' have been developed, where strategies such as role play, pictures and games have been used to help teach perspective-taking (Tse, Strulovitch, Tagalakis, Meng & Fombonne, 2007). The study conducted by these authors indicated positive results, but its non-controlled pre-/post-test design had obvious limitations. The results of controlled studies of the efficacy of such interventions have not always demonstrated meaningful outcomes in 'real life' situations (Ozonoff & Miller, 1995; Reichow, Steiner & Volkmar, 2013) and reviews of the literature have concluded an insufficient and incomplete evidence base (White, Keonig & Scahill, 2007; Bellini, Peters, Benner & Hopf, 2007). The breadth of skills

that are addressed in social skills programmes and disparity between programmes are just a couple of the factors that make evaluation more difficult (Rao, Beidel & Murray, 2008).

Smith (2001) cites a number of issues pertaining to the implementation of social skill programmes in mainstream school settings including difficulties in gathering together a sensitive and relevant group for which the intervention will be pertinent to all, the typically high oral language load of such interventions, and the lack of generalisation of new skills to the everyday life experiences of pupils. As such, interventions which are focused at a more individual level are often deemed to be more feasible.

Although individual differences exist, researchers largely concur that, for individuals with ASC, processing difficulties are particularly salient when the stimuli presented is of a transient or non-spatial nature (Baranek, 2002; Hutchins & Prelock, 2006). This assertion, combined with the findings of research that has suggested that the use of visual strategies and cues (such as written prompts and graphics) are helpful in improving social communication and behaviour (e.g. Mesibov & Howley, 2003; Quill, 1997; Theimann and Goldstein, 2001) has particularly been drawn upon in the development of personalised interventions aimed at improving the social communication and interaction skills of pupils with ASC, such as those outlined below.

2.4.4 Social Stories

One popular intervention is Social Stories (Gray, 1994a) which is now commonly used in UK schools (Jones 2002; Smith, 2001). Social Stories are short, personalised stories written about a particular social situation which a pupil finds challenging. A specific format is used which provides information about where and when the situation occurs, who is involved, what usually occurs and why. Included within this is 'perspective' information, which describes the reactions and feelings of others, and directive information which describes what the child should try to say or do in that situation.

The aim is to improve an individual's understanding of events and expectations by providing the information that they may be currently missing, in order to enable them to produce more effective responses in future situations (Gray & Garand, 1993; Gray, 2000). As such, Social Stories aim to target the social communication and interaction difficulties inherent in autism. Gray (2004) claims that social stories can be used to address a variety of issues including helping children to follow rules and routines, increasing appropriate behaviours, or decreasing inappropriate behaviours.

There has been a fairly significant amount of interest in evaluating the effectiveness of Social Stories and a number of reviews of the literature have been conducted. A likely reason for this interest is the apparent simplicity of the approach and the fact that it is an inexpensive and non-time demanding intervention. However, the outcomes of research thus far have generally been inconclusive. Ali & Frederickson (2006) argued that there was a sufficient evidence base to suggest that the intervention has promise and Karkhaneh et al, (2010), focusing only on controlled trials, found significant benefits for a variety of outcomes related to social interaction. Another review has concluded that, whilst Social Stories appeared to have low overall effectiveness, they were more effective when addressing inappropriate behaviours than when teaching social skills (Kokina & Kern, 2010). Reynhout and Carter's (2011) meta-analysis similarly concluded that Social Stories are of questionable or mild efficacy. Consistent amongst all reviews has been the assertion that further, more robust research is needed.

An alternative to Social Stories is Comic Strip Conversations (CSCs), also devised by Gray (1994b), which shall be described in the following section.

2.5 Comic Strip Conversations

2.5.1 What are Comic Strip Conversations?

An alternative to Social Stories is Comic Strip Conversations (Gray, 1994b). This lesser-known intervention is similar to a Social Story as both techniques involve visual systems designed to support understanding of social situations

and encourage more appropriate social behaviours. One key difference is that it is intended that the individual with ASC is an active participant in the construction of a CSC, alongside another person. Through the creation of CSCs, not only can the child learn important social information, so too can the adult learn about the perspectives of the child (Gray, 1998).

CSCs are used to review a situation taken directly from an individual's life and to discuss alternatives to behaviour which had proven to be unbeneficial in that situation (Pierson & Glaeser, 2005). Visual representations are used to demonstrate some of the more abstract aspects of social interaction, such as recognising other people's thoughts and feelings. Simple stick figures and other drawings are used to symbolise people and objects, and symbols such as speech bubbles and thought bubbles are used. As already outlined in section 2.3.1, thought bubbles seem to make some sort of intuitive sense to children, enabling them to overcome difficulties in the mental representation of others' thoughts and beliefs (Kerr & Durkin, 2004). Colour can be used to identify the emotional content of a statement, thought or question. For example, green can be used for happy words and red for angry words. Colours can be gradually introduced as appropriate over the course of several 'conversations'. Through such visual techniques, it is hoped that abstract concepts, such as thoughts and feelings, can be made more concrete and thus easier to understand.

2.5.2   Structure of CSCs

Gray (1994b) recommends that the activity is introduced in a way which demonstrates that drawing whilst talking is an acceptable way to communicate. Ideally, the child or young person takes the lead and is encouraged to write, draw and talk the majority of the time.

The child and adult sit next to each other, with their joint attention focused on the work area. Boxes may be drawn to help order the sequence of events. A representative location symbol is drawn in the upper left hand corner of the work area to identify the location of the topic of the conversation.

The adult guides the 'conversation' through asking questions which aim to gather information about the situation. These questions are expected to include at least some of the following, used as appropriate:

- Where are you?

- Who else is here?

- What are you doing?

- What happened? What did others do?

- What did you say?

- What did others say?

- What did you think/feel when you said that?

- What did others think/feel when they said/did that?

The adult can share their perspective with the child as and when appropriate, for example, if they are having difficulty answering a question. The goal is to achieve a balance between gathering insights into the child's perspective whilst sharing accurate social information. Given the nature of the social difficulties in ASC, questions regarding the thoughts and feelings of others are likely to be the most difficult to answer. If a child's answer demonstrates errors in determining what others are thinking, another idea is introduced without discrediting their response.

After summarising the CSC, the final step is to identify possible solutions to the situation. If the child cannot do this independently, a solution is suggested and recorded, before asking if they can think of any others. Sometimes, a number of possible solutions are identified and the pros and cons of each may be discussed. These discussions can also be recorded through drawings to provide visual support. Solutions deemed unfeasible are removed, and the child is left with an 'action plan' for what to do the next time the situation occurs. Some examples of CSCs (taken from the present research) can be viewed in Appendix 2.

2.5.3 How May CSCs Help Individuals with ASC?

The approach, through its discussion and representation of mental states, draws on the concept of theory of mind in its aim to support the 'mind reading' skills of an individual with ASC, with the ultimate aim being to support and ameliorate social interaction and communication difficulties (Gray, 1998).

Gray (1998) also proposes that the creation of CSCs can support central coherence. Through the creation of a CSC, contextual information is described which can aid in establishing relationships between relevant cues and defining meaningful responses. In other words, 'central coherence' information is provided through the means of pictures and writing which may help those individuals who are otherwise less able to make those links independently.

As with other story-based interventions (such as Social Stories), CSCs aim to provide an individual with accurate social information which can support them in knowing how to respond appropriately in a given situation. Rigidity of thought and action is one of the triad of impairments originally identified by Wing & Gould (1979) and there is a tendency for individuals with autism to rigidly adhere to rules once they have been internalised (Scattone, Wilczynski, Edwards & Rabian, 2002). As CSCs can provide concrete, social information, and include the formation of a visually recorded action plan, this could aid behaviour change through providing individuals with a representation or 'rule' of what they should do in a specific situation.

2.5.4 The Evidence Base of CSCs

Gray (1998) suggests that CSCs may be used in conjunction with Social Stories or independently. Indeed, there exists a small body of research which has evaluated the impact of joint Social Story/CSC interventions with positive results being indicated (Hutchins & Prelock 2006, 2008, 2012). However, compared to Social Stories, the efficacy of CSCs in their own right has attracted far less attention from researchers and it is only in relatively recent years that literature about the intervention has gone beyond simply proposing strategies for effective CSC construction (Hutchins & Prelock, 2006). In the

remainder of this chapter, a systematic review of the research on the efficacy of CSCs as an intervention for individuals with ASC is presented with the aim of clarifying and summarising the current knowledge base and identifying areas for further investigation.

2.6   Systematic Literature Review

The systematic literature review will report on the currently available evidence relating to the effectiveness of CSCs for individuals with autism.

Systematic reviews are a particular way of sourcing and synthesising research evidence in a given area and are closely linked to the evidence-based movement (Robson, 2011). Within an educational context, the evidence-based movement stipulates that decisions about approaches and methods adopted by practitioners should be based upon systematic knowledge of intervention outcomes (Dunsmuir, Brown, Iyadurai & Monsen, 2009).

Systematic reviews can be a means of contributing to knowledge bases about what does or does not work, and can identify where further research may be needed (Petticrew & Roberts, 2008). They aim to find as much as possible of the research relevant to particular questions, and use explicit methods to identify what can reliably be said on the basis of the research found (EPPI-Centre, no date).

2.6.1   Review Question

What is the effectiveness of CSCs for improving outcomes in individuals with ASC?

2.6.2   Systematic Search

A systematic search was undertaken in order to identify the available research related to CSCs. A keyword search was conducted using two databases relevant to the fields of study which are of interest in the present review. The databases searched were PsycInfo and ERIC. The search terms used were:

'comic*' AND (autis* OR asperger*)

Due to the small number of studies that were found through the database search, a keyword search was also conducted on Google Scholar with the aim of capturing any articles that were not available on the databases. As a final search strategy, the reference lists of the retrieved and relevant articles were scanned in order to increase confidence that all the presently available research had been located.

2.6.3   Inclusion criteria

In order to ensure that the articles obtained reflected research that was relevant and up-to-date, a set of inclusion criteria were applied (see Table 2.1).

Inclusion criteria
• At least one participant had a diagnosis, or working diagnosis, of an ASC • If other non-ASC participants were included, the effects of the intervention on the participant(s) with ASC could be examined individually • An intervention involving CSCs had been conducted • If CSCs were part of an intervention 'package', then data regarding the impact of CSCs could be isolated from other aspects of the intervention • The effectiveness of this intervention had been measured by way of outcomes for the participant(s) • The research was published in a peer reviewed journal. • The research was published between 1990 and 2015. • The research was in the English language.

Table 2.1: Inclusion criteria applied in the systematic literature search

2.6.4 Search Results

Following the systematic search, seven articles were identified as meeting the inclusion criteria (see Table 2.2). A total of 20 articles were excluded for reasons such as them not being a research paper or not being relevant to the review question. Three articles were found which evaluated CSCs and Social Stories as a joint intervention but these were excluded as the relative contribution of the CSC aspect of the intervention had not been isolated, therefore it was deemed that such papers could not elucidate the efficacy of CSCs. A further article was found which had evaluated the effectiveness of CSCs for students who did not have ASC, therefore this was excluded due to it not being focused on the population of interest in the present research. See Appendix 3 for further details of the excluded articles.

Source	Papers included in review
ERIC	Glaeser, Pierson & Fritschmann (2003) Pierson & Glaeser (2007) Robinson (2008) Rogers & Myles (2001)
PsycInfo	*1 additional article sourced:* Ahmed-Husain & Dunsmuir (2014)
Google Scholar	*2 additional articles sourced:* Vivian, Hutchins & Prelock (2012) Lewandowski, Hutchins, Prelock & Murray-Close (2014)
Reference lists of relevant articles	No additional articles sourced

Table 2.2: List of research articles sourced through the systematic review which met the inclusion criteria.

A key facts table outlining details of all seven studies can be found in Appendix 4.

2.6.5 Research Quality and Relevance Appraisal

The methodological quality and relevance of each study included in the review was assessed using the Weight of Evidence (WoE) Framework (Gough, 2007), summarised in Table 2.3. Judgements were made of the 'WoE' that each study could provide for answering the review question (high, medium, or low).

WoE A: Methodological quality	WoE B: Methodological appropriateness	WoE C: Relevance of evidence	WoE D: Overall weight of evidence
Generic judgement on the quality of execution of the study	Appropriateness of the research design for answering the review question	Relevance of the focus of the study to the review question	Overall assessment of the extent to which a study contributes evidence to answer the review question

Table 2.3: Weight of Evidence framework (Gough, 2007)

WoE A is a generic and non-review-specific judgement about the coherence and integrity of the evidence in its own terms and can be assessed using generally accepted criteria for evaluating the quality of this type of evidence (Gough, 2007). To inform WoE A for the non-experimental case study designs, Barker, Pistrang & Elliott's (2002) list of features that improve the credibility of this type of evidence were considered. These include:

- The use of systematic and quantitative data
- Multiple assessments of change over time
- Multiple cases
- Change in previously chronic or stable problems
- Immediate or marked effects following the intervention

For the studies which used SCEDs, the 'Quality Indicators Within Single-Subject Research' criteria outlined by Horner et al (2005) and Kratochwill et al's (2013) SCED design standards were considered when making judgements. These include core elements such as:

- Operational definitions of the dependent variable(s)

- Sufficient measurement occasions during each phase (ideally, at least five)

- Sufficiently consistent pattern of responding documented during baseline

- Reliability measures for the dependent variable (e.g. inter-observer agreement)

- Active manipulation of the independent variable, with experimental control established through the use of at least three data series to allow for rival hypotheses for positive results (e.g. maturation, history) to be discounted

- Efforts to enhance external validity, such as through replication of effects across participants, behaviours or contexts.

WoE B is a review-specific question about the appropriateness of that form of evidence for answering the review question (Gough, 2007). In this case, that amounted to how adequately the design could answer the efficacy-related review question. Factors such as the nature of the design (e.g. level of experimental control) and the use of appropriate and reliable measures contributed to this judgement. A 'high' rating was given if the design and analysis used was deemed entirely appropriate for answering the review question A 'medium' rating therefore represents a reasonable level of appropriateness of the design and analysis for answering the review question, and a 'low' rating represents a judgement that the design and analysis used is unclear or inappropriate for answering the review question.

WoE C is a review-specific judgement about the relevance of the focus of the evidence for the review question, for example the type of sample, the type of evidence gathering or the nature of the context (Gough, 2007). A judgement of 'high' was made if the study was of direct relevance to the review question. A judgment of 'medium' represents that the study was deemed relevant to the review question to some extent. A 'low' judgement indicates that the study was only deemed to be indirectly or insufficiently relevant to the research question.

When making an overall judgement (WoE D), all individual judgements were considered and an average judgement decided on. Further information regarding the judgement of the weightings can be found in the following paragraph and also in section 2.6.6.1.

Table 2.4 demonstrates that the majority of the studies were rated as providing an overall 'low' weight of evidence for the review question. This was mainly the result of these studies lacking empirical control and a lack of compensatory mechanisms that could have increased the coherence and integrity of study outcomes (as reflected in the 'A' and 'B' ratings). Only one study was rated as providing a 'high' weight of evidence (Ahmed-Husain & Dunsmuir, 2014) and this was due to its research design demonstrating better control for extraneous variables, the inclusion of multiple participants and the use of statistical, as well as visual, analyses to provide a more reliable interpretation of the results.

Study	WoE A	WoE B	WoE C	WoE D
Rogers & Myles (2001)	Low	Low	Medium	Low
Glaeser, Pierson & Fritschmann (2003)	Low	Low	High	Low
Pierson & Glaeser (2007)	Medium	Low	High	Medium
Robinson (2008)	Low	Low	Low	Low
Vivian, Hutchins & Prelock (2012)	Low	Low	Medium	Low
Ahmed-Husain & Dunsmuir (2014)	High	High	High	High
Lewandowski, Hutchins, Prelock & Murray-Close (2014)	Medium	Medium	High	Medium

Table 2.4: 'Weight of Evidence' ratings given to the studies included in the systematic review

2.6.6   Synthesis of Research

What follows is a narrative synthesis of the included literature which provides an overview of the key themes and factors related to the reviewed research. In order to remain focused on the review question of interest, if a study included

participants without a diagnosis of ASC (in addition to participants with ASC) then the results for these participants are not included in the reporting.

2.6.6.1 Design

One key theme that emerged from the present literature review was that of the research design used. Due to the paucity of empirical research available at present with regards to the efficacy of CSCs, research using non-experimental designs were included in the present review, although the caution that needs to be applied when interpreting the outcomes of such studies will be considered.

Case studies:

Of the seven studies included in this review, five utilised variations of case study design (Rogers & Myles, 2001; Glaeser et al, 2003; Pierson & Glaeser, 2007; Robinson, 2008; Vivian et al, 2012). All the case studies reported positive results to varying extents. There are obvious limitations to the findings produced from case studies that lack any form of experimental control as the threats to internal validity are significant. There is also poor external validity as generalisation of results is not possible.

Barker, Pistrang & Elliott (2002) identify the following features that improve the credibility of evidence from non-experimental case studies:

- The use of systematic and quantitative data

- Multiple assessments of change over time

- Multiple cases

- Change in previously chronic or stable problems

- Immediate or marked effects following the intervention

Of the five case studies, only Rogers & Myles (2001) included clear quantitative data but it was deemed that this could not be defined as 'systematic'. Pierson & Glaeser (2007) presented percentage reduction figures as part of their findings, but it appeared that these had been extrapolated from qualitative data collected by staff. Pierson & Glaeser's (2007) study reported

on three cases which enhances the generalisibilty of the results, and Robinson's (2008) study involved two participants. All studies reported a change in previous problems, but a lack of information made it difficult to assess the 'stability' of these problems or how long they had been present for. The final point was similarly difficult to assess in some studies, although the results from Pierson & Glaeser (2007) and Vivian et al (2012) indicate a marked effect following the intervention as judged by school staff and parents respectively.

Single-Case Experimental Designs

The remaining two studies (Ahmed-Husain & Dunsmuir, 2014; Lewandowski et al, 2014) used single-case experimental designs (SCEDs). These provide more robust evidence due to the incorporation of a baseline phase which acts as a control (for a particular individual) and repeated measures being taken over time.

An ABA multiple-baseline across behaviours design was used by Ahmed-Husain & Dunsmuir (2014). In such a design, the intervention is introduced at different times for different behaviours - if there is a corresponding change in the condition to which the intervention is applied, but no change in the other condition at that time, it provides a strong case for inferring causality. In an ABA design, repeated measures continue to be taken following the withdrawal of the intervention - if a return to baseline performance is observed then there is a reasonably strong case for inferring a causal link. However, issues arise when the nature of the intervention means that an effect of withdrawal is less likely, for example, if something has been learned that will not easily be 'unlearned'. Given that CSCs ultimately aim to help address theory of mind impairments, effects that are maintained even after the intervention ends might be expected.

Lewandowski et al (2014) adopted an ABACA design in which, following a standard ABA sequence, a new 'form' of the intervention (which involved the sibling of the participant with ASC also completing the intervention) was introduced and subsequently withdrawn.

As with cases studies, there are issues of external validity in SCEDs. The multiple participants (n=8) in the Ahmed-Husain & Dunsmuir (2014) study helped to increase external validity. Given the heterogeneous nature of the ASC population, designs which aim to demonstrate external validity can be problematic.

2.6.6.2 Participant Characteristics

In total across the seven included studies, 17 participants (11 male, six female) who had a diagnosis of an ASC received a CSC intervention. The sample size ranged from n=1 to n=8. The stated ages of the participants ranged from five to 14 years old. Robinson (2008) did not clarify the age of the participants in her study, but as it took place in a further education setting it can be assumed they were at least 16 years of age.

The participants were variously described as having a diagnosis of mild or high-functioning autism (Glaeser et at, 2003; Pierson & Glaeser, 2007), ASD or ASC (Robinson, 2008; Vivian et al, 2012; Ahmed-Husain & Dunsmuir, 2014) or Asperger syndrome (Rogers & Myles, 2001; Ahmed-Husain & Dunsmuir, 2014; Lewandowski et al, 2014).

One study (Ahmed-Husain & Dunsmuir, 2014) outlined the results of a standard cognitive assessment for each participant. This indicated a very wide range of cognitive ability amongst the participants, ranging from the 1st percentile to the 95th percentile. The participant in Vivian et al (2012) underwent baseline assessments of expressive and receptive vocabulary and non-verbal intelligence and scored in the average range or above average range on these measures. This same participant scored in the low average range on measures of theory of mind and social skills but scored very high on a problem behaviour scale.

2.6.6.3 Target Behaviours

The studies reviewed involved a wide variety of target behaviours. These are outlined in Table 2.5. The majority of studies focused on just one or two

behaviours for each participant. The study by Vivian et al (2012) stands out due to the varied foci of the CSCs that were created

Study	Behaviour(s) targeted through CSC intervention
Ahmed-Husain & Dunsmuir (2014)	Across 8 participants: Making eye contact; fiddling with objects in lessons; banging and tapping; making inappropriate comments to others; shutting eyes and putting head on desk during lessons; asking for help less often; initiating conversations
Rogers & Myles (2001)	Needing numerous redirections and being late to P.E lesson after lunch
Glaeser et al (2003)	Conflicts with adults and peers
Lewandowski et al (2014)	Conflict with brother
Pierson & Glaeser (2007)	Across 3 participants: Appropriate use of hands and feet on the playground; social greetings using eye contact and appropriate voice volume; accepting responsibility for inappropriate actions and apologising to peers. In addition to the target behaviours, this study was also interested in analysing levels of loneliness and social satisfaction, which it was considered would change if improvements were seen in target behaviours.
Robinson (2008)	Communicative intent and interaction relevant to the content of tutorials
Vivian et al (2012)	Crying and shouting at bedtime; getting 'time-out' at school; talking back to teachers; being called names; responding appropriately to requests; responding appropriately when offered unwanted food; telling brother to do something he should not; talking to others about your feelings; hitting friends; being grumpy.

Table 2.5: Behaviours targeted through the use of CSCs in the studies included in the systematic review

2.6.6.4   CSC Intervention Procedures

Format:

The amount and quality of the information provided about the intervention procedures actually used varied between articles. A couple of the studies specified that Gray's (1994b) guidelines were followed in the creation of CSCs. In the remaining studies where this wasn't explicitly stated, there is evidence that at least some of the key features of the procedure proposed by Gray (1994b) were utilised. Some articles included images of completed CSCs which demonstrated the use of some of the key features of the approach (Glaeser et al, 2003; Vivian et al, 2012; Lewandowski et al, 2014). Only two studies indicated that intervention fidelity checks were conducted (Lewandowski et al, 2014; Vivian et al, 2012). Both these studies reported very good fidelity rates, thus enhancing the validity of their findings.

Some researchers incorporated additions or amendments to the standard procedures outlined by Gray (1994b). In order to ascertain the relative impacts of different forms of 'action plan', the participants in Ahmed-Husain & Dunsmuir's (2014) study completed one CSC which included a 'visual' action plan (as would typically happen if following standard procedures) and another CSC which included an 'auditory' plan. Two studies (Vivian et al, 2012; Lewandowski et al, 2014) also included affirmative CSCs. The potential that this could have had on motivation levels and acceptance of the intervention should be acknowledged as a factor which may have impacted on the results for their participants.

Level of Participant Involvement:

With regards to the level of participant involvement in the creation of CSCs, this could differ in terms of how much of the drawing and writing a participant completed, and how many of the 'solutions' were generated by the participant as opposed to the adult they were working with.

The Rogers & Myles (2001) study appears to indicate that the teacher took responsibility for the creation of the CSCs, which does not tally with the joint

approach to creation which is encouraged by Gray (1994b). However, very limited procedural details are provided in this article making it difficult to know if this was the intended procedure, or if the participant expressed an unwillingness to contribute to the drawing and writing. Robinson's (2008) study does not make enough details clear to establish the relative input of the participants and their tutors.

The remaining studies indicate varying levels of participation in the drawing, writing and solution-generating processes. However, CSCs are, by their nature, a dynamic intervention. Therefore, whilst guidelines exist, response to the intervention and the degree of participation will be expected to vary between individuals (Hutchins & Prelock, 2006).

Setting and Interventionists:

Five studies took place in an educational setting. Of these, three were in a mainstream education setting (Rogers & Myles, 2001; Glaeser et al, 2003; Ahmed-Husain & Dunsmuir, 2014), one was in a special education setting (Pierson & Glaeser, 2007) and one was in a further education setting (Robinson, 2008). In all these studies, the CSC intervention was implemented by staff members. The remaining two studies took place in the home setting. The intervention was implemented by the parents of the participant in one of these studies (Vivian et al, 2012) but by a researcher in the other (Lewandowski et al, 2014).

Frequency and Duration:

There were notable differences in the frequency of CSC creation and the overall duration of the intervention period amongst the included studies. For the studies in which frequency of implementation information was available, the number of CSCs created over the course of the intervention periods ranged from one (Ahmed-Husain & Dunsmuir, 2014) to 24 (Lewandowski et al, 2014). In some studies, the regular review of previously created CSCs was a key part of the process (Pierson & Glaeser, 2007; Ahmed-Husain & Dunsmuir, 2014). In two studies (Rogers & Myles, 2001; Glaeser et al, 2003), the duration of the intervention is unclear, although the latter study makes reference to 'the

first few months' of using CSCs, implying a relatively extended implementation period. Robinson (2008) indicates a 3-month period of using CSCs in tutorials. The other studies specified intervention periods of four to six weeks (Ahmed-Husain & Dunsmuir, 2014), six weeks (Pierson & Glaeser, 2007; Vivian et al, 2012) and 89 days followed by another 75 days after an interim withdrawal phase (Lewandowski et al, 2014).

None of the studies aimed to explore the potential differential effects of frequency of CSC use or intervention duration. It has been suggested that the number of repetitions needed to establish positive outcomes is likely to vary across individuals, targets and contexts (Gray, 1998; Vivian et al, 2012). In addition, it has been advised that a CSC should be repeated several times before abandoning it as adjustments or additions may be required to achieve the desired outcome (Vivian et al, 2012).

Outcome Measures:

The principal and, in many cases only, outcome measure used in the majority of the studies was observational data. The nature of this varied amongst studies. See Table 2.6 for an overview of the measures used and the person(s) responsible for taking these measures.

In the studies involving observational measures, the observations were either carried out by the individual(s) who were also responsible for delivering the intervention or the researcher. Therefore, the potential for observational bias is present, especially in the studies that used qualitative observational methods (Glaeser et al, 2003; Pierson & Glaeser, 2007; Robinson, 2008). In one study where a teacher and researcher conducted joint observations (Ahmed-Husain & Dunsmuir, 2014), the inter-observer agreement was calculated to be between 87% - 95% which provides increased confidence in the reliability of their data.

Two studies (Lewandowski et al, 2014; Vivian et al, 2012) included subjective measures of parental opinions regarding their child's behaviour which could have been prone to demand characteristics.

Study	Outcome Measure(s) Used	Measures Completed By:
Ahmed-Husain & Dunsmuir (2014)	Interval sampling observation	Researcher and teacher simultaneously
Rogers & Myles (2001)	Event sampling	Teacher
Glaeser, Pierson & Fritschmann (2003)	Qualitative observation / anecdotal data	Staff
Lewandowski, Hutchins, Prelock & Murray-Close (2014)	Daily diary, including a rating scale Theory of Mind Inventory	Parent Parent
Pierson & Glaeser (2007)	Qualitative observation and anecdotal records	Teacher and Teaching Paraprofessional
Robinson (2008)	Observation within an ethnographic framework	Researcher
Vivian, Hutchins & Prelock (2012)	Theory of Mind Inventory Post-intervention interview	Parents

Table 2.6: Outcome measures included in the studies and who was responsible for taking the measures

2.6.6.5 Results

All the articles included in this review reported positive results for at least some, if not all, participants. In this section, further details of the results of each study shall be provided.

Results from case studies:

Of the research involving case study designs, the reservations about which have already been stated earlier in this review and should be borne in mind when considering the validity of the results, have all reported positive findings.

Rogers & Myles (2001) concluded that CSCs had been more effective than a Social Story intervention for the participant in their study. However, it is not

clear how they came to this conclusion as a decrease in the target behaviours had already been recorded during the prior implementation of the Social Story. Their conclusion may have arisen through reports that the student was enjoying using CSCs and had begun to independently request their use at home and school. It should be noted that in this study, although the outcome measures reported related to the number of verbal redirections needed and the number of minutes late to P.E class after lunchtime, the actual content of the participant's CSCs were said to be related to social incidents which occurred during lunchtime. Whilst it was hypothesised that the target behaviours were the result of these lunchtime issues, the fact that the outcome measures did not reflect the actual CSC content makes the evaluation of CSCs in this study quite distinct.

Glaeser et al (2003) asserted that the participant in their study had shown 'great progress in dealing with conflicts with peers and adults' (p. 17) over the first few months of using CSCs and, as she became more independent in writing the CSCs herself, she became better able to respond more appropriately to her classmates in a natural and age-appropriate manner.

Pierson & Glaeser (2007) reported 'significant changes' in target behaviours for all three participants, with between 50% - 75% improvement from baseline. This study also cites evidence for decreased levels of loneliness and greater social satisfaction, such as fewer loneliness verbalisations, increased talkativeness with peers, more smiles and a greater desire to participate with peers inside and outside the classroom.

The parents of the participant in Vivian et al's (2012) study stated that the use of CSCs had 'definitely helped' facilitate more appropriate behaviours. In addition, her parents had also highlighted the feasibility of the intervention, describing it as 'quick' and 'simple'. Whilst this was not a focus of the review question, it was deemed an important point to note as issues of feasibility and ease of use are of real importance when implementing interventions in applied settings. This study also reported the scores achieved on a Theory of Mind

Inventory[1] (Hutchins, Prelock & Bonazinga, 2012) which indicated an increase from the 17th to the 50th percentile between pre- and post-intervention, although this finding cannot be causally attributed to the intervention.

Robinson (2008) found that the two participants demonstrated increased levels of communicative intent and interaction relevant to the content of tutorials during 'CSC tutorials' compared to during a typical tutorial. This study is unique amongst the others in this review as CSCs were used as a way to structure thoughts and conversation during a tutorial situation with the aim of improving communication in that same situation. So, in effect, CSCs were used to improve social functioning 'there and then' as opposed to using it as a tool to review a past situation and consider alternative ways of behaving in similar future situations.

Results from single-case experimental designs:

For the reasons outlined earlier in this review, the results of the studies employing a SCED are considered to be more robust. The two studies utilising SCED designs (Ahmed-Husain & Dunsmuir, 2014; Lewandowski et al, 2014) also reported positive findings. Ahmed-Husain & Dunsmuir found, overall, the CSC intervention to be effective for seven out of eight participants. Four out of eight participants responded well to both CSCs ('visual' action plan and 'auditory' action plan) and three out of eight participants responded well to one of the CSCs. These results were found to be maintained in the medium term.

These researchers also found that participants' verbal and visual skills matched the type of action plan used for the most successful CSC, suggesting a link between cognitive strengths and the type of CSC that was more effective. These results challenge the assumption that visual interventions always work best for individuals with ASC. However, for one participant in this study neither CSC was effective. The authors considered that this may have been a reflection of the complex nature of the target behaviours (responding

---

[1] This inventory is a psychometrically evaluated tool which consists of 48 statements assessing a range of theory of mind abilities which are answered by a child's caregivers.

appropriately to other students and asking for help from the teaching assistant less often). It was also this participant who had the lowest cognitive ability score (1st percentile), indicating that such a factor may be relevant when considering for whom CSCs may be most effective.

As two of the participants in the Ahmed-Husain & Dunsmuir (2014) study appeared to demonstrate delayed intervention effects, the authors concluded that the effectiveness of CSCs could be dependent on the length of time they are implemented, as well as how often they are reviewed.

Lewandowski et al (2014), in their ABACA design, found no intervention effect between phase A and B, but a positive effect was found between the second A phase and the C phase which was maintained in the medium term. This was seen as suggestive of an impact when there was further opportunity to complete CSCs, thus indicating the potential utility of a longer intervention duration. However, as the outcome measure of this study was subjective parental ratings of behaviour, this makes the results less reliable than those found by Ahmed-Husain & Dunsmuir (2014). Lewandowski et al (2014) also used the Theory of Mind Inventory (Hutchins et al, 2012) as an additional pre- and post-intervention measure, which indicated a slight improvement (from 4th to 8th percentile) between assessment sessions but this could not be causally attributed to the intervention.

In SCEDs, achieving stability in the baseline is of importance as this will demonstrate reliability in the target behaviour, making subsequent analysis of any changes during the intervention phase much easier (Barlow, Nock & Hersen, 2009). It should be noted that the baselines for the participants in the Ahmed-Husain & Dunsmuir (2014) study were, in a number of cases, quite unstable, as was the baseline for the participant in Lewandowski et al's (2014) study. This means that intervention effects inferred through visual analysis should be treated with caution. However, both studies conducted additional statistical analyses to interpret the data.

2.6.7 Summary and Conclusion of the Systematic Review

The systematic review is limited by the small number of studies available and the small numbers of participants involved, thus limiting generalisation. In addition, all but one of the studies originated in the U.S.A. Differences between the education system there and in the UK may further make generalisation difficult.

The included studies reported positive outcomes for all (or, in one case, most) of the participants with ASC in terms of improvements in particular behaviours. Whilst most studies have focused on decreasing undesirable behaviours, there are a few examples of CSCs being used to increase desirable behaviours (such as making eye contact and initiating conversation). The nature of the behaviours targeted has been varied and has included some (e.g. rocking on chair, fiddling with objects) that do not immediately strike as being 'social' in nature. This is perhaps an early indication of the potential scope of CSCs in terms of the behaviours they could be used to address.

These preliminary positive results have been found across participants of different ages and in different settings. In some cases, positive effects beyond the immediate target behaviours of interest have been indicated (e.g. increases in scores on a theory of mind inventory), although methodological issues mean these cannot be confidently attributed to the intervention.

At present, case studies constitute the majority of the currently available evidence for the efficacy of CSCs. Whilst it is deemed that such evidence contributes some understanding to this topic, questions regarding the reliability and validity of this type of evidence cannot be ignored. The emerging use of SCEDs to evaluate the efficacy of CSCs has provided additional, more robust, evidence for the potential of the approach. However, a lack of external validity is still an issue in SCEDs. Studies that include multiple participants (e.g. Ahmed-Husain & Dunsmuir, 2014) help to increase external validity. However, given the heterogeneous nature of children and young people with ASC, establishing external validity will always be problematic.

A few further examples of SCED research exist which incorporate CSCs and Social Stories as a combined intervention, with efficacy being assessed accordingly. Such research was not included in the present systematic review due to it not being solely focused on the specific intervention of interest and the impossibility of evaluating the relative contribution of CSCs to outcomes. However, the existence of such research demonstrates an increasing interest in empirical evaluation of CSCs as an intervention to target the difficulties seen in individuals with ASC.

It is clear that study design needs to be carefully considered in future research. Whilst randomised controlled trials are generally considered to be the 'gold standard', the relatively low prevalence of students with ASC make such designs unfeasible as they require a large number of participants in order to achieve adequate statistical power (Ahmed-Husain & Dunsmuir, 2014). Furthermore, the appropriateness of such designs could be queried due to the heterogeneous nature of the ASC population.

It is not yet clear for which children CSCs are likely to be the most effective. It appears that the majority of participants in the studies reviewed could be described as 'high functioning' but, in general, limited descriptions of participants' abilities across potential domains of relevance mean that, at present, it is not possible to produce a profile of the children who may most benefit from CSCs. A small number of researchers have, however, begun to hypothesise about the potential impact of factors such as cognitive ability (Ahmed-Husain & Dunsmuir, 2014) and verbal mental age (Vivian et al, 2012) on levels of responsiveness to the intervention.

Inconsistencies regarding intervention frequency, duration and procedures in the studies reviewed further complicate the process of developing a coherent understanding of the conditions under which CSCs may be most beneficial. Therefore, a further direction for future research could be to investigate how such factors may impact on the effectiveness of the intervention.

Despite the preliminary promise of CSCs as an intervention to address the social communication and interaction difficulties in children and young people

with ASC, the empirical evidence is still at an early stage. It is apparent that more research is needed before it can be confidently (or not) recommended as an intervention for children and young people with ASC.

2.7 Rationale for the Present Research

As previously outlined, there exists at present only a small body of research into the efficacy of CSCs with very little of it being experimental in design. The need for more objective and experimental evaluations of CSCs has been recognised by various authors (Gray, 1998; Hutchins & Prelock, 2006; Ahmed-Husain & Dunsmuir, 2014). The present research, therefore, aimed to build on the limited empirical evidence that currently exists.

The researcher's literature search indicated that no experimental study had yet been published that assessesd the efficacy of CSCs with primary-aged pupils in a mainstream school setting, so this became the focus of the present research. The research also aimed to address questions regarding the frequency with which to implement a CSC intervention, as previous research had shown significant inconsistencies in this area. Such a question was deemed to be worth exploring as, in the applied school setting in which the research was conducted, time is never abundantly available. If CSCs can be demonstrated to be effective even at low levels of implementation frequency then this can 'free up' valuable time to spend on other objectives. Alternatively, it is possible that increased frequency of implementation would result in opportunities to apply the approach in response to multiple situations, thus enhancing the effects of the intervention due the increased opportunities for generalisation.

From a more pragmatic stance, a further rationale for the present research was that CSCs was an intervention that some of the EPs and staff from the Autism Outreach Team (AOT) within the authority in which the researcher was on placement recommended to schools on a regular basis. An opportunity to have the intervention evaluated within the local context was welcomed. A further point to note is that, at present, most of the current research comes from the USA, with the exception of one study that took place in the UK (Ahmed-

Husain & Dunsmuir, 2014). Therefore, further research within a UK context is needed.

In terms of the wider rationale for the present research, now that many individuals with ASC are educated in mainstream settings, it is very important that staff in such settings have access to, and the knowledge to implement, practical and inexpensive interventions to help support the needs of their pupils (Ali & Frederickson, 2006). CSCs, which intend to meet these criteria, offer one potential means to do just that.

2.8 Research Questions and Hypotheses

Table 2.7 displays the research questions addressed along with their experimental and null hypotheses.

Research Question	Hypotheses
Research Question 1: *1a:* Do CSCs have a positive impact on the target behaviours of primary-aged pupils with ASC?	*Experimental hypothesis:* The CSC intervention will lead to improvements in target behaviours (i.e. a decrease in frequency). *Null hypothesis:* The CSC intervention will have no effect on target behaviours.
1b: Are the findings of the repeated measures reflected in perceived change in target behaviours from the perspective of school staff?	*Experimental hypothesis:* The CSC intervention will lead to a positive impact on staff perceptions of the target behaviours. *Null hypothesis:* There will be no change in staff perceptions of the target behaviours.
Research Question 2: Does the level of improvement in target behaviours vary in relation to the frequency with which the CSC intervention is implemented?	*Experimental hypothesis:* There will be a differential effect on target behaviours depending on intervention frequency - greater improvement will be seen when CSCs are produced weekly. *Null hypothesis:* There will be no differential effect on target behaviours depending on intervention frequency.

Table 2.7: Research questions and their experimental and null hypotheses

Chapter 3 -   Methodology

3.1   Introduction to Methodology

Following an introduction to issues related to evidence-based practice, this chapter provides an account of the methodology used in the present research. The rationale behind the methodological decisions made is outlined and matters relating to the reliability and validity of the research are discussed.

3.2   Real World Research

Real world research refers to research conducted in applied fields, such as education, which focuses on issues of direct relevance to people's lives (Robson, 2011). The present research was conducted in school settings, the nature of which can make it difficult to manipulate independent variables and measure outcomes with precision (Stoiber & Waas, 2002). As Robson (2011) identifies, the contexts of 'real world' research are often highly complex and challenge the researcher to say something sensible about 'messy' situations, placing findings within the real context of a phenomena as opposed to within the context of a laboratory.

3.2.1   Evaluation Research

Much real-world research takes the form of evaluation (Robson, 2011). Robson (2011) differentiates between formative evaluation and summative evaluation. Summative evaluation aims to assess the effectiveness of an intervention and therefore focuses on outcomes. In contrast, the purpose of formative evaluation is to explore how an intervention may be effective, therefore the focus is on processes and mechanisms. The present research aimed to provide a summative evaluation of CSCs which can be compared with the existing evidence base.

3.3   Evidence-Based Practice

Within the profession of educational psychology, the role of EPs as 'scientist practitioners' can be enhanced through identifying evidence-based interventions (Stoiber & Waas, 2002). Repeated calls have been made for more rigorous evaluation of interventions at both group and individual level (Frederickson & Miller, 2008). A key goal in school based interventions is to identify what works in improving outcomes, and a variety of research methods can be utilised in gaining knowledge of intervention effects (Stoiber & Waas, 2002). The overall aim of the present research was to contribute to evidence-based practice through the evaluation of an intervention in school settings.

Originating from the medical profession, Roth & Fonagy (1996) developed a hierarchy of evidence which determines the quality (i.e. reliability and validity) of the various types of research design that are available to draw upon in professional practice. These are outlined in Table 3.1

Rank	Type of Evidence
1	Several systematic reviews of Randomised Controlled Trials
2	Systematic review of Randomised Controlled Trials
3	Randomised Controlled Trials
4	Quasi-experimental trials
5	Case-control and cohort studies
6	Expert consensus opinion
7	Individual opinion

Table 3.1: The hierarchy of research evidence (adapted from Frederickson, 2002)

As demonstrated in Table 3.1, RCTs dominate the upper ranks of the hierarchy. However, some have queried the appropriateness of this hierarchy within educational research, as statistical averaging (as occurs in RCTs) obscures the individual responses that might lead to a better understanding of the conditions

under which, and with whom, interventions are most likely to be effective (Neef, 2009). Thus, whilst RCTs can answer the efficacy question of 'Can it work?' (when studied within maximally controlled environments), the effectiveness question of 'Does it work?' requires research which replicates the circumstances of everyday practice (Harrington, 2001).

Frederickson (2002) suggests that the type of research evidence specified in levels four and five of the hierarchy have the advantage of being compatible with the day-to-day practice of EPs although concedes that, due to the lack of key controls in such research, establishing evidence of efficacy is more difficult. Ultimately, the research approach adopted depends on the research question(s) being asked - one should aim for the best available evidence of the most appropriate type (Frederickson, 2002). The research questions asked and the ways in which attempts are made to answer them are informed by the philosophical assumptions of the researcher, as shall be discussed in the following section.

3.4   Theoretical Paradigms and Philosophical Assumptions

A paradigm is a way of looking at the world, composed of philosophical assumptions that guide thinking and action. Researchers need an understanding of these paradigms and philosophical assumptions, and need to locate their own research within a paradigm, in order to successfully plan and carry out their research (Mertens, 2010). Underpinning these paradigms are sets of assumptions of an ontological, epistemological and methodological nature.

3.4.1   Ontology, Epistemology and Methodology

Ontological assumptions are concerned with the nature of reality or of a phenomenon (Cohen, Manion & Morrison, 2011). A researcher's ontological standpoint is determined by whether they believe that social reality is an objective phenomenon (i.e. external to an individual) or a subjective phenomenon (i.e. the result of an individual's cognition) (Cohen et al, 2011). Ontology is the beginning point for all research and epistemological and methodological decisions will follow from this (Grix, 2002).

Epistemology is concerned with the nature of knowledge and the relation between the knower and the would-be known (Guba & Lincoln, 2005). It addresses questions about what knowledge is out there to be known and how such knowledge can be acquired (Cohen et al, 2011).

Informed by ontological and epistemological assumptions, methodology is concerned with the approach to systematic enquiry taken by a researcher (Mertens, 2010).

3.4.2   Dominant Paradigms in Educational and Psychological Research

Cohen et al (2011) describe two dominant paradigms within educational and psychological research - positivism and constructivism. Table 3.2 outlines the ontological, epistemological and methodological assumptions that underpin these two major paradigms, which have traditionally been viewed as being incompatible with one another (Fox, 2002).

Philosophical Assumptions	Positivism	Constructivism
Ontology	There is one reality that can be known within a specified level of probability	There are multiple, socially constructed realities.
Epistemology	Objectivity is of importance; the researcher manipulates and observes in an objective manner (mirroring the way the natural world is studied).	There is an interactive link between researcher and participants; findings are 'created' (i.e. knowledge is subjective); values are made explicit
Methodology	Primarily quantitative; interventionist and de-contextualised	Primarily qualitative; hermeneutical, dialectical; contextual factors considered

Table 3.2: Ontological, epistemological and methodological assumptions associated with the positivist and constructivist paradigms (adapted from Mertens, 2010).

3.4.3 Post-Positivism

As a result of the tension between positivism and constructivism, and questions about the applicability of the positivist paradigm within 'real world' research (Robson, 2011), an alternative paradigm has become prominent within applied research: post-positivism. Post-positivism adapts the approaches taken in the natural sciences to apply them to social science research (Robson, 2011). Post-positivists concur that a reality does exist and seek to identify causal relationships, but appreciate that this reality can only be known imperfectly (Mertens, 2010).

Whilst this paradigm strives for objectivity, with the researcher remaining neutral in order to prevent bias from entering the research, it is accepted that the theories, hypotheses, knowledge and values of a researcher can bear an influence. In addition, post-positivists argue that a number of the assumptions necessary for rigorous application of scientific method are difficult, if not impossible, to achieve within many educational and psychological research studies (Robson, 2011).

3.4.4 Perspective of the Current Research

In this research, the researcher worked within a post-positivist paradigm and this informed the design decisions made and the methods used. Within educational provision, the concept of evidence-based practice has been traditionally set within the post-positivist paradigm (Fox, 2002). This research aimed to make a contribution to the drive for evidence-based practice through investigating causal relationships using quantitative data within a single-case experimental design (SCED).

In keeping with the post-positivist paradigm, the researcher aimed to achieve a level of experimental control. However, the existence of extraneous factors within the real-life context in which the research was conducted were acknowledged, and therefore the potential limitations in relation to interpreting and generalising findings were recognised (Robson, 2011).

3.5   Research Design

This research employed a series of multiple-baseline across behaviours SCEDs. A separate experiment was conducted for each of the five participants involved. For reasons that shall be outlined in section 3.5.3 an A-B SCED (rather than multiple-baseline) was used for one of the five participants.

3.5.1   Single-Case Experimental Designs

SCEDs offer a way of identifying causal relationships and establishing evidence-based practices. In contrast to group comparison approaches, SCEDs focus on the individual and 'are organised to provide fine-grained, time-series analysis of change in a dependent variable(s) across systematic introduction or manipulations of an independent variable' (Horner et al, 2005, p.172). A level of experimental rigour is established through the 'case' serving as their own control via a comparison of performance prior to, and during and/or after, an intervention. Table 3.3 outlines the key characteristics of SCEDs.

Single-case research is popular among researchers and practitioners in education and psychology. Horner et al (2005) propose that the use of SCEDs is particularly appropriate when the aim is to consider the performance of a specific individual under a given set of conditions. The aim of the present research was to consider the impact of the intervention on specific, target behaviours that were unique to that individual. Horner et al (2005) also propose that research questions that can be appropriately addressed with SCEDS include those which focus on the relative effects of two or more independent variable manipulations on the dependent variable(s), and this research additionally focused on the relative effects on behaviour following different levels of exposure to the intervention.

The focus at an individual level does lead to questions about the generalisation of findings from SCED research (Barlow et al, 2009). This will be discussed further in section 3.10.2.

Feature	Definition	Purpose
Continuous assessment	Measures taken on multiple occasions over time during both baseline and intervention phases.	Provides the basic information on which data evaluation and identification of intervention effects depend.
Baseline assessment	Assessment over a period of time prior to implementing the intervention.	Describes current performance and allows prediction of how performance is likely to continue in the immediate future.
Stability of performance in baseline	Stable performance is one in which there is relatively little variation over time.	Permits projections of performance to the immediate future and allows evaluation of the impact of an intervention.
Use of different phases	Periods of time in which a specific condition (baseline or intervention) is implemented and data collected.	To test whether performance continued in the predicted pattern from a prior phase or changed as the intervention was implemented. Inferences about intervention effects are drawn from the pattern of data across phases.

Table 3.3: The key characteristics of SCEDs as identified by Kazdin (2003)

3.5.2   Types of SCED

There exists a range of design variations that can be considered when designing a SCED, differing in levels of complexity and the extent to which they can answer cause and effect questions. Table 3.4 outlines the key characteristics of common SCEDs, along with their relative advantages and disadvantages.

Kratochwill et al (2013) have outlined a set of criteria for SCEDs that meet what they term 'evidence standards'. Included criteria are as follows:

- the independent variable is systematically manipulated
- there are at least three attempts to demonstrate an intervention effect at three different points in time
- a phase must contain a minimum of three data points

With regards to the final criteria, it should be noted that some authors have argued for a minimum of five data points (Horner et al, 2005).

Design	Design Features	Advantages of Design	Disadvantages of Design
AB	Involves a baseline phase (A), followed by a further sequence of observations (B). The effectiveness of the intervention is shown by difference in the two phases of observation.	Can be useful in an initial pilot study and has high applicability to professional practice	Low internal validity due to difficulties in ascribing causal explanation
ABA	As AB, but a third phase is added which reverts to pre-intervention baseline condition.	Higher internal validity than AB designs, as the reversal can demonstrate the effect of the IV	Ethical concerns in withdrawing an intervention
ABAB	Involves the addition of a second intervention phase (B) to an ABA design.	Further enhances internal validity	Ethical concerns in withdrawing an intervention (although it is reinstated)
Multiple-Baseline	1. Across settings – dependent variable is measured in two or more situations. Change is made from a baseline condition (A) to the intervention (B) at different times in the different settings 2. Across behaviours - two or more behaviours are measured, with changes between A and B phases made at different times for the different behaviours. 3. Across participants - two or more participants are measured, with changes made at different times for the different participants.	Greater internal validity than AB designs - if there are changes in subject to which the intervention is applied, but not to other cases at that time, there is a stronger case for arguing causal relationships.	Interventions which show generalisation across target behaviours cannot be evaluated within a single case.

Table 3.4: Key features of different SCEDs and their advantages and disadvantages (adapted from Robson, 2011)

3.5.3 Present Research Design and Rationale

For four out of the five participants, this research employed a multiple-baseline across behaviours SCED in which the researcher investigated the effects of a CSC intervention on two separate target behaviours. In order to address the second research question, the frequency of intervention implementation differed across the two behaviours. This enabled the researcher to explore the relative impact of intervention frequency on outcomes. A multiple-baseline design was chosen as it allowed for two attempts at demonstrating an intervention effect at two different times, thus enhancing internal validity.

For one participant, following a couple of weeks of baseline data collection, it appeared that one of the target behaviours identified during the initial stage of the research had all but extinguished, meaning there was little reason to implement an intervention to target it. Therefore, the decision was made to focus on a single behaviour for this participant as there were no obvious replacement behaviours which would meet the criteria required by the study in terms of frequency and suitability to be addressed through CSCs. As such, it was considered that an A-B SCED offered the most appropriate way forward for this participant, albeit with acknowledgement of the limitations of such a design when it comes to identifying causal relationships.

3.5.4 Consideration of Alternative SCEDs

An A-B design was rejected at the initial design stage due to the lack of internal validity inherent in such a design. An A-B-A design was rejected as it was not deemed appropriate to withdraw a potentially beneficial intervention. In addition, it was considered that a reversal effect may not be demonstrated as the intervention may lead to behavioural changes that could not be readily 'undone' simply by discontinuing it. This reason was also applicable to the rejection of an A-B-A-B design, with time constraints also making such a design unfeasible.

3.5.5 Other Research Designs Considered

A key focus of this research was to evaluate the effectiveness of the CSC intervention through identifying possible causal relationships. With an aim such as this, the researcher was led to consider a more traditional group design, such as an RCT (Robson, 2011). However, given that the research was to be conducted in the real-life context of schools with what was (identified at the planning stage) going to be a small number of participants, this made the use of an RCT design inappropriate. In addition, the heterogeneous nature of the population of interest would limit the effectiveness of attempts at randomisation and matching procedures used in RCTs and alternative quasi-experimental designs (Odum et al, 2003).

3.6 Stakeholders

In planning the present research, a number of stakeholders were considered:

- The University of Nottingham
- The Educational Psychology Service in which the researcher was on placement at the time of the research
- The Autism Outreach team (AOT) based in the local authority in which the researcher was on placement
- The schools who participated in the research, the participants and their parents
- The researcher, a doctoral student and trainee educational psychologist

3.6.1 The relevance of the CSC intervention to stakeholders

The researcher had identified during her time on placement that CSCs were being recommended by a number of EPs as an intervention for a range of children (both with and without ASC) who were presenting with social and behavioural issues of concern to schools. It was also known that the AOT delivered training on CSCs to schools and that the intervention was commonly recommended by them. As such, both the EPs and the AOT would be interested in the results of the study.

3.6.2   Stakeholders and Time Scale

The time scale for the research was determined by the researcher who was completing a doctorate over three years. Participant recruitment began during the summer term of the researcher's second year of study. In discussion with staff at the educational psychology service in which the researcher was on placement, it was agreed that data collection would be completed during the autumn term of the researcher's third year of study. Due to some unforeseen circumstances, the data collection for three of the participants continued into the spring term.

3.7   Participants

3.7.1   Sampling Strategy

A purposive sampling strategy was used in which participants were identified in order to satisfy the specific requirements of the research (Robson, 2011). This approach was taken because it fitted the research design and allowed the researcher to select participants from a particular population (primary-aged pupils with a diagnosis of ASC) that would satisfy the needs of the research question and the research rationale.

3.7.2   Inclusion Criteria

In order to ensure that the participants chosen were appropriate in terms of being able to meet the needs of the research questions, a number of inclusion criteria were developed. These were as follows:

1.   The pupil has a diagnosis of ASC (to include Asperger's)
2.   The pupil attends a mainstream primary school
3.   The pupil is displaying (or not displaying) specific behaviours that are deemed to be appropriate for a CSC intervention
4.   These behaviours occur frequently enough for a sufficient number of examples to be observed within time-limited observation periods

5.   The pupil has sufficient receptive and expressive language skills to be able to access the intervention

6.   Participants must not already be receiving an intervention specifically aimed at targeting the behaviours wished to be targeted through CSCs.

3.7.2.1   Rationale for Inclusion Criteria 1, 2 and 3

The first three criteria relate to the purpose of the research which was to investigate the effectiveness of CSCs in addressing the target behaviours of interest for primary-aged pupils with a diagnosis of autism who were being educated within a mainstream school. The justification for these criteria has previously been outlined in section 2.7. Behaviours were deemed to be appropriate for a CSC intervention if they were viewed as having a detrimental impact on a pupil's learning and/or their social relationships (with staff and/or peers) within school as it was important that they presented as 'socially valid' intervention targets (Horner et al, 2005).

3.7.2.2   Rationale for Inclusion Criteria 4

It was necessary that the frequency of the target behaviours was sufficiently high in order to ensure that they could be adequately captured by the regular observational measure required within the SCED and thus provide an adequate means of identifying any changes in frequency as a result of the CSC intervention. The specific nature of the observations were decided on once sufficient information had been gathered about each participant, as will be explained in 3.8.2.

3.7.2.3   Rationale for Inclusion Criteria 5

As CSCs necessarily involve a certain level of discussion, it was important to determine that participants' language skills were of a level that would not preclude them from accessing such discussions. Language skills were not formally assessed as it was felt that a one-off assessment with an unfamiliar adult may not provide an accurate picture of their abilities. In addition, the time demands of carrying out comprehensive language assessments with each

participant would have been impractical for the researcher. Questions regarding each participant's language skills were rather addressed during the information gathering process involving staff.

3.7.2.4   Rationale for Inclusion Criteria 6

It was appreciated by the researcher that it was likely that, given their diagnosis of ASC, potential participants would be receiving some form of intervention or adaptation to their learning environments in order to better to meet their needs within school. It was considered that stating that no interventions could currently be being undertaken would be impractical and would severely limit the number of participants available. As such, it was rather specified that the participants must not already be receiving an intervention specifically aimed at targeting the same behaviours that would be addressed through the CSC intervention (as this would clearly have implications when determining possible causal relationships).

3.7.3   Participant Recruitment

The initial step in the recruitment process involved the researcher sending a letter about the research (see Appendix 5) to the Headteachers of mainstream primary schools within the patch of schools in which the researcher and their placement supervisor were working at that time. Follow-up phone calls were made to schools where necessary.

This process ultimately resulted in expressions of interest from five schools relating to six different pupils[2]. It was considered that this would be a sufficient number to help increase the external validity of the findings, whilst at the same time being practical in terms of the scope of the project and the time available to the researcher.

Informed parental consent was sought through providing parents with written information about the research (see Appendix 6) and they were also offered an

---

[2] One participant was subsequently lost during the baseline phase of the research due to staffing issues

opportunity to meet directly with the researcher to discuss the research and ask any questions. All parents provided written consent.

Semi-structured interviews were then held with school staff. Semi-structured interviews consist of pre-determined questions but the order of these and the wording used can be modified as deemed appropriate, and particular questions can be omitted or added depending on the responses of the interviewee (Robson, 2011). They generally comprise open-ended questions which can, if being used as a data collection measure, be more difficult to analyse than closed questions. However, the interviews were being used to gather sufficient information about the participants, rather than as a unit of measurement relating to the effectiveness of the intervention, therefore this point was not considered a hindrance. See Appendix 7 for a copy of the interview proforma.

A key purpose of these meetings was to identify potential behaviours that could be targeted through the intervention. It also allowed the researcher to gather more general information about the participant, such as the current provision available to them in school, to enable a fuller description of the participants. In most cases, the staff involved in these interviews were the school's special educational needs co-ordinator and a teaching assistant who regularly worked with the pupil and would be the member of staff delivering the intervention.

3.7.4   Participant consent

Participant consent was sought through providing participants with a child-friendly information sheet (see Appendix 8). This information sheet was shared with the pupils by a member of staff who they knew well and, in all cases, was the member of staff who would be delivering the intervention. In the cases involving the two youngest participants, staff supported their understanding of the information sheet through drawings. Participants were asked if they would be happy to take part in the research and, if so, they were asked to write their name on the information sheet. There was a more formal participant consent form available for completion should the staff member feel that this was an appropriate consent tool to use. All participants provided consent.

3.7.4.1   Staff consent

Informed consent was also sought from the staff members as they were highly involved in the research through delivering the intervention and, in most cases, the data collection. Staff members were provided with an information sheet which outlined the respective responsibilities of the researcher and the staff member (see Appendix 9). The information sheets were personalised for each member of staff to reflect differences in data collection procedures. All staff members provided consent.

Further information regarding the ethical issues surrounding consent is discussed in section 3.11.

3.7.5   Description of Participants

The participants were five pupils in years one to four. All participants were male. Details about each participant are provided in Table 3.5. All names used are pseudonyms to protect the anonymity of the participants.

Participant	Age at research onset	Year group	Statement of SEN?	Diagnosis	Key areas of difficulty as identified by staff	Support arrangements within school
Jack	6 years and 10 months	2	No	ASC	• Social interaction skills such as turn-taking and recognising/appreciating other people's personal space • Low-level classroom disruption such as calling out and getting out of his seat regularly.	• No one-to-one support • No interventions were being implemented at the time of the onset of data collection. • During the course of the research, Jack begun receiving weekly 'positive play' sessions, targeting skills such as listening to instructions, sharing and making appropriate eye contact.
Daniel	6 years and 6 months	2	Yes	ASC	• Social communication skills, especially engaging in two-way conversations • Obsessive-type behaviours • Poor attention and listening • Low academic attainment	• One-to-one support in lessons for 20 hours per week • Use of autism-friendly resources such as visual timetables and visual prompt cards. • Regular small group activities in Maths and Literacy
Robert	8 years and 4 months	4	No	ASC ADHD (not medicated)	• Difficulties in peer relationships • Is easily led by others • Angry outbursts, resulting in swearing and pushing peers.	• One-to-one support in the classroom for 15 hours per week • Short session with teaching assistant each morning to discuss

					• Lack of independent learning skills	'making good choices' with regards to friendship issues. • Own copy of instructions for tasks are provided
Owen	5 years and 9 months	1	No	ASC (working diagnosis)	• Difficulties in peer relationships, especially in relation to co-operating with others and wanting to do things 'his way', sometimes leading to shouting or physical obstruction • Low academic attainment	• No one-to-one support, but receives regular support within a small group during Maths and Literacy • Weekly motor skills intervention group • Speech and language targets are addressed in school
Gareth	8 years and 6 months	4	No	ASC	• Poor attention and listening skills and lack of independent learning. • Not following instructions • Fiddling with own and others' belongings • Low academic attainment and an apparent disinterest in much of the curriculum content in school	• One-to-one support for 15 hours per week • Individual workstation used for some lessons • Now and next boards and other visual prompts used within classroom.

Table 3.5: Summary of key participant details

3.7.6 Target Behaviours

Target behaviours were initially identified through the process of semi-structured interviews. To ensure that the behaviours were socially valid targets for intervention (Horner et al, 2005), the behaviours chosen were those deemed by staff to be having a negative impact on the participant's learning and/or social relationships. These conversations were supported by data from an exploratory narrative observation of each participant carried out by the researcher. Further informal consultation with school staff then enabled the clarification of the target behaviours for each participant, including the development of operational definitions of these. Table 3.6 displays the behaviours selected for each participant.

Participant	Target Behaviour 1 (TB1)	Target Behaviour 2 (TB2)
Jack	*Not waiting his turn to speak:* Jack demonstrates behaviours such as calling out inappropriately and interrupting a member of staff when they are taking to somebody else. Not to include instances of Jack calling out at a time when a significant number of other children are doing the same.	*Touching/squeezing peers' faces or bodies:* Jack demonstrates behaviours such as putting his arms around another person and squeezing them or putting his hands on their face/head and squeezing, or other instances of uninvited touching. Not to include instances of appropriately timed contact (e.g. reciprocating a hug off a peer) or accidental contact.
Daniel	*Asking time-related enquiries:* Daniel asks time-related questions such as, 'Is it morning/afternoon?' 'Is it 12 o'clock?' 'What time is it now?'	*Waving finger in other people's faces:* Daniel waves one or more of his fingers close to somebody else's face. Not to included instances of Daniel waving his fingers in front of his own face only.

Robert	*Not starting tasks independently:* Robert does not start a task when instructed to - e.g. he sits and does nothing or something unrelated to task, or he waits for an adult to come over Not to include instances of Robert immediately and independently reasonably seeking adult support following an instruction to start a task (e.g. putting his hand up to say he does not understand).	N/A
Owen	*Calling out 'No!':* Owen says or shouts 'No!' in response to a statement, instruction, action, or at any time when such a response is inappropriate. Not to include instances of Owen responding with 'No' in an appropriate way (e.g. as a appropriate response to a closed question).	*Interfering with peers' use of toys/objects during free-choice time:* Owen snatches or attempts to snatch items being used by his peers, knocks things out of peers' hands as they are using them or damages/destroys something his peers have made (e.g. lego models). Not to include instances of accidental damage (e.g. accidently stepping on something in his way)
Gareth	*Calling out in class:* Gareth calls out during lessons rather than waiting his turn (e.g. calling out an answer for a question directed to a peer; calling out comments while the teacher or a peer is speaking).	*Touching or taking items belonging to others:* Gareth touches or takes items that do not belong to him when he has not had permission to do so (e.g. touching peers' belongings such as pencil cases or school bags/taking someone else's

	Not to include instances of Gareth calling out at a time when a significant number of other children are doing the same.	pen). Not to include instances of touching or taking items when permission has been given to do so.

Table 3.6: Target behaviours of each participant with examples of what does or does not constitute an instance of the behaviour.

3.8 Measures

Within this section, the independent and dependent variables within the research will be outlined. This is followed by a description of each measure used in relation to the dependent variables and issues of validity and reliability related to these.

3.8.1 Research Variables

The independent variable was the CSC intervention. There were two versions of this as will be explained in section 3.9.7. The dependent variables were:

1. The frequency of target behaviours

2. Staff perceptions of the level of challenge or disruption posed by the behaviours, and perceptions of the impact of the behaviours on the participant's learning and/or social relationships.

3.8.2 Measure of Dependent Variable 1: Structured Behaviour Observation

In SCEDs, the dependent variable is typically an observable behaviour (Horner et al, 2005). A major advantage of observation is its directness and it is often considered the appropriate technique for accessing 'real life' in the real world (Robson, 2011). Structured observation is systematic and enables the collection of numerical data which, in turn, facilitates the calculating of frequencies, patterns and trends (Cohen et al, 2011).

For four of the five participants, two target behaviours were observed through repeated structured observations. For one participant (Robert), only one behaviour was subject to structured observation.

Frequency of observations:

Repeated measurement is critical for comparing performance between phases and sufficient assessment occasions are necessary in order to establish an overall pattern. Some suggest that three data points per phase is the minimum acceptable (Kratochwill et al, 2013), whilst others recommend a minimum of five (Horner et al, 2005). In order to achieve at least this higher figure, it was planned for two observations to be conducted per week. In one case (Owen), the observations for TB2 were only able to be conducted once per week as this behaviour occurred during a particular context ('free choice' time) which was only timetabled on a Friday afternoon.

Structured Observations:

In structured observations, a coding scheme needs to be developed which contains predetermined categories for recording what is observed (Robson, 2011). In developing the coding schemes used in this research, the researcher aimed to adhere to Robson's (2011) considerations to make it as straightforward and reliable as possible, as outlined below:

- Focused: Looking only at the selected target behaviours
- Objective: Requiring little inference from the observer
- Explicitly defined: Use of a sufficiently detailed operational definition of the behaviour, with examples of what does and does not fall within the category
- Exhaustive: Covering all possibilities of what the behaviour may look like in order for it to be possible to make a coding when the behaviour occurs
- Easy to record: For example, just ticking a box.

In order to address practical issues, in addition to enhancing ecological validity, the precise nature of the structured observations varied between

participants due to various factors. Firstly, the typical frequency of the target behaviours helped to determine the length of the observational period and the contexts in which the observations were conducted. A second factor involved issues of staff availability. In some instances, a lack of staff availability resulted in the researcher having to take responsibility for the observations. In the cases where staff members were conducting the observations, it was important that the schedules they used could be completed with ease during their time with the participant without having a detrimental impact on their day-to-day work supporting the participant and, in some cases, other pupils as well. Table 3.7 outlines the nature of the observational schedule for each participant.

In the cases where staff members were conducting the observations, they received a training session in the use of the observational protocol in order to familiarise themselves with how to complete it. Then followed a short trial period in which the schedules could be tested for ease of use and to ensure that the definitions of the target behaviours were clear. If needed, conferencing between the researcher and staff member was used to further clarify the target behaviours in order to reduce ambiguity and improve the reliability of the observational system. If no issues arose during this trial period, the data collected was included in the baseline data set.

In most cases, observations took place across a range of lessons at different times of the day to help ensure that a variety of situations were observed where the behaviour had been identified as an issue by staff, and to capture as realistic a picture as possible about the general frequency of the behaviour. In some cases, target behaviours predominantly occurred in particular contexts, namely Jack and Owen's TB2, therefore observations took place in the appropriate settings.

Event sampling was chosen as the most appropriate observational method in most cases as it is a useful method for finding out frequencies of behaviours and simply requires a tally mark to be entered each time the target behaviour is observed (Cohen et al, 2011). This made it a convenient method for staff

members who needed to be able to continue their activities with the participant during the observational period.

Interval sampling was used in Jack's observational schedule, which was completed by the researcher. Interval sampling is a highly structured observation approach and this method was chosen as the researcher was able to apply the stringent time-keeping necessary. Such an approach would not have been possible in the cases in which staff members were collecting the data. Examples of the interval and event sample observation schedules used can be viewed in Appendix 10.

Participant	Observational Schedule for TB1	Observational Schedule TB2	Observations conducted by:
Jack	Interval sample over a 30-minute period in the classroom (25 seconds observation / 5 seconds recording).	Interval sample over a 10-minute period during breaktime (25 seconds observation / 5 seconds recording)	Researcher
Daniel	Event sample over a 30-minute period in lessons.	Event sample over a 30-minute period in lessons.	Teaching Assistant
Robert	Event sample across the school morning.	N/A	Teaching Assistant
Owen	Event sample over a 30-minute period in the classroom.	Event sample over a 30-minute period during 'free choice' time (once per week)	Teaching Assistant for TB1 Researcher for TB2
Gareth	Event sample over a 45-minute period in the classroom.	Event sample over a 45-minute period in the classroom.	Teaching Assistant

Table 3.7: Observational schedule for each participant

3.8.3 Measure of Dependent Variable 2: Pre- and Post-Intervention Questionnaires

In addition to the repeated measures, further measures were obtained at two points in time - firstly, prior to the start of the baseline phase and, secondly, at the end of the intervention phase. The collection of pre- and -post intervention data within SCEDs has been documented by a number of researchers (e.g. Nock, Michel & Photos, 2007). Whilst it is important to acknowledge the limitations of data collected at only two points in time (discussed further in section 3.10.3.3), this data was supplementary to that obtained through the repeated measures and was included for data triangulation purposes.

A brief self-completion pre- and post-intervention questionnaire was designed by the researcher and these were completed by staff members. As the focus of this research was on the specific target behaviours of each participant, there were no pre-existing questionnaires available that were appropriate for answering the first research question. It was considered that a self-report questionnaire was an appropriate method for gathering the required information as the respondents were able to complete the questionnaire at a convenient time. It was hoped that this would allow them to reflect on their responses to a greater extent than they may have if an alternative procedure had been used, such as an interview. It was also considered that this method could reduce the influence of factors relating to the presence of the researcher (Robson, 2011).

The pre-intervention questionnaire comprised rating scales to assess staff perceptions of how challenging or disruptive they found the target behaviour to be and how much of an impact they perceived it to have on the participant's learning and/or social relationships. The post-intervention questionnaire included these same questions along with an additional rating scale to assess how effective the staff member felt the CSC intervention had been in addressing the target behaviour.

Rating scales are widely used in research as they combine the opportunity for a flexible response with the ability to determine quantitative analysis (Cohen et

al, 2011). One key limitation of rating scales is that the numbers can have different meanings for different respondents (Cohen et al, 2011). However, as this research was focused on individual cases, rather than aggregating and averaging results across participants, the impact of this limitation was considered to be reduced.

As rating scales do not allow respondents to provide additional comments about the issue under investigation, the use of open-ended questions can be useful in smaller scale research to gather information that would otherwise not be captured (Cohen et al, 2011). For this reason, the post-intervention questionnaire also included an opportunity for staff to provide additional comments relating to their perceptions of the participant's response to the intervention.

See Appendix 11 for copies of the pre- and post-intervention questionnaires.

3.9   Procedure

3.9.1   Piloting

The researcher had originally considered carrying out a small scale pilot of the research but ultimately did not. In some cases, aspects of a piece of research can make piloting difficult to set up and also less important (Robson, 2011). The key reasons for not piloting the research were:

- Each individual case would utilise different observational schedules (reflecting the nature of the target behaviours and who would be observing it) therefore there was not a 'standard' schedule to pilot.

- The researcher had previously delivered CSC training outside of the research context and was able, as a result of feedback following this, to better ensure that training was comprehensive and supportive to the staff who would be involved in implementing the intervention.

- Difficulties and delays experienced in participant recruitment meant there were insufficient numbers to assign a participant to a pilot study.

Despite the lack of a formal pilot study, it should be noted that all observers involved in the research had the opportunity to trial the observational measures developed for each participant to allow for any issues to be identified and resolved through conferencing if necessary.

3.9.2 Staff training

In order to ensure that the intervention was delivered in a consistent manner and following clear guidelines, the researcher herself attended a staff training session delivered by an experienced educational psychologist, the content of which was based on Gray's (1994b) guidance. Subsequently, the researcher delivered a staff training session (outside of the research context) in order to become further familiarised with the intervention and to gather feedback regarding how well the training equipped staff with the confidence to deliver the intervention themselves. The researcher also liaised with a member of staff from the AOT who provided the researcher with a copy of their version of CSC training, which was very similar in content to that delivered by the educational psychologist.

From all this information, the researcher generated a CSC training protocol and manual. All staff designated to deliver the intervention attended a standardised training session delivered by the researcher. The session lasted approximately one hour. This included a general introduction to the theory behind CSCs, information about how to create a CSC, and an opportunity to create a 'practise' CSC. Staff retained a copy of the training materials and a one-page 'quick guide' to creating CSCs to assist them in delivering the intervention. They were also supplied with a copy of Gray's (1994b) book about CSCs. A copy of the training materials can be found in Appendix 12.

3.9.3 Collection of Pre-Intervention Measures

Prior to the commencement of each participant's baseline phase, staff completed the pre-intervention questionnaire, which was then collected in person by the researcher.

3.9.4 Baseline Phase

As is expected with multiple-baseline designs, the length of the baseline phase varied between participants and between behaviours. Baseline phases lasted between two and a half and five weeks. In some cases, there were breaks in the baseline data collection due to school holidays.

The stability of the baseline is important in SCEDs as this will demonstrate reliability in the target behaviour, making subsequent analysis of any changes during the intervention phase much easier (Barlow et al, 2009). Whilst the researcher acknowledged that, ideally, baseline phases should be extended until sufficient stability is achieved, timescales and ethical considerations placed restrictions on their length in this research.

3.9.5 Intervention Phase

When planning the length of the intervention phase, the researcher considered previous research into CSCs. As noted in the systematic literature review, previous research has demonstrated differences in the overall duration of the intervention period. With CSCs not being an 'off the shelf' intervention programme, there is no stated intervention period. With this is mind, the researcher planned for an intervention period which it was considered would allow for:

- each participant to have a consistent and sufficient amount of exposure to the intervention
- a sufficient number of data points to be collected
- completion of the research within the restricted time scales

The length of the intervention phase ranged from three to five weeks. As with the baseline phase, the length of the intervention phase is expected to vary in multiple-baseline designs. During the intervention phase, data continued to be collected (as far as possible) in accordance with the schedules outlined in Table 3.7.

3.9.6 Timescales of Baseline and Intervention Data Collection

It had been originally planned for the data collection to be completed by the end of the autumn term. However, delays in participant recruitment and commencement of data collection resulted in some data collection continuing until the February half-term. Difficulties in conducting consistent weekly observations of Owen's TB2 were experienced owing to timetable changes preceding the Christmas holiday. This resulted in missing data and led to the continuation of data collection for this behaviour after the Christmas holiday, even though the data collection for his TB1 had finished. This matter is further addressed in section 4.5.4.4.

3.9.7 Two Intervention Conditions

In order to answer Research Question 2, there were two alternative CSC intervention conditions. In one condition, a single CSC (termed 'Comic Strip Conversation Single' - CSCS) was created to address one of the two target behaviours. In the second condition, weekly CSCs (termed 'Comic Strip Conversation Weekly' - CSCW) were created to address the second target behaviour. The purpose of this was to ascertain whether or not a higher intervention frequency made a difference to the effectiveness of the intervention and, therefore, whether or not there may be potential value in repeating the intervention regularly. Further information regarding the rationale for this additional research question has previously been outlined in section 2.7.

Due to timescales, it was anticipated that between three and four CSCs would be completed in the CSCW condition. Four was considered to be the maximum desirable as it was considered that completing more than this

relating to the same target behaviour may have resulted in lowered motivation for the participant. The two target behaviours of each participant (with the exception of Robert, for whom only one behaviour was targeted) were randomly allocated to either the CSCS or CSCW condition. The order of presentation of the two conditions was also randomly assigned in order to protect against possible influences related to intervention novelty.

3.9.8   Implementation of CSC Intervention

As recommended by Gray (1994b), the staff members completed an introductory CSC with each participant towards the end of the baseline phase on a topic completely unrelated to the target behaviours of interest in this research. The purpose of this was to familiarise the participants and staff members with the process and highlight any operational issues prior to completing the first CSC aimed at addressing a target behaviour.

The CSC intervention was implemented by a member of school staff, following the guidelines they had received during the training sessions. In all cases, the staff member was a teaching assistant. In three cases (Daniel, Gareth and Robert), the teaching assistant delivering the intervention was also the participant's key worker who supported them across part of the school day. In the other two cases (Jack and Owen), the participants did not have a key worker, but the teaching assistant worked within their classes and was therefore familiar with, and to, the participant.

Intervention schedules were created to support staff to implement the intervention with the planned frequency for each behaviour. Intervention sessions generally took place on the same day each week, however there was some variability due to changes to school timetabling and staff availability. Each CSC created was unique as target behaviours manifested themselves differently between participants.

3.9.8.1 Reviews of CSCs

In addition to the main intervention sessions, regular review sessions were planned to be conducted at least three times per week. The review sessions lasted approximately five to ten minutes and involved the staff member reminding the participant about previously completed CSCs. The main purpose of this was to ensure that the CSCs were not simply created and then put away and forgotten about. Staff were provided with a set of prompts that they could use as appropriate to help structure these review sessions. These included:

- Has the topic of the CSC occurred since it was last discussed?
- Has the pupil tried out their 'action plan'? If so, how did this go?
- If necessary, remind pupils of their 'action plans'.

In addition, if the participant had further or new ideas for their 'action plans', they were permitted to add these to their CSC. It was considered important that the participants maintained a sense of ownership of their CSCs as the intervention is one which aims to involve the pupil as an active participant in the process. This level of flexibility, in addition to incorporating review sessions, has been encouraged by Vivian et al (2012) who suggest that 'adjustments or additional dialogue may be needed to secure desirable outcomes' (p. 40).

Throughout the course of the research, the researcher kept in weekly contact with participating school staff to discuss any arising operational issues and monitor progress.

3.9.9 Intervention diaries

Throughout the intervention phase of the research, staff members completed an 'intervention diary' (see Appendix 13). The purpose of this diary was two-fold. Firstly, it allowed the researcher to ascertain how much exposure each participant had had to the intervention. Although intervention schedules were agreed upon at the outset, the researcher acknowledged that circumstances could arise in which these schedules were not adhered to, thus resulting in

under-exposure to the intervention. It was deemed that such information could potentially be of relevance when considering the outcomes for different participants (e.g. the number of CSCs and review sessions actually completed).

The second purpose of the intervention diary was to allow staff to record any information that they deemed pertinent to the implementation of the intervention, for example, any amendments made to CSCs or comments about the participant's level of engagement. Again, it was considered that such information could be of relevance when considering the results. Whilst it was intended that the intervention be delivered following standard guidelines as far as possible, the researcher appreciated that the nature and extent of the participants' contributions during intervention sessions was likely to vary due to the heterogeneous nature of the participants in terms of factors such as age, ability level and motivation.

3.9.10 Intervention Fidelity

Intervention fidelity refers to the degree to which specified procedures within an intervention are implemented. A lack of intervention fidelity will compromise the internal validity of outcome studies (Dane & Schneider, 1998).

The researcher considered assessing intervention fidelity through direct observation of intervention sessions. However, this idea was disregarded on two grounds. Firstly, it was considered that the presence of the researcher could potentially have a detrimental (e.g. stressing) effect on the participant and staff member. Secondly, as the intervention necessarily creates tangible artefacts (i.e. the completed CSCs), it was deemed that examination of these would be able to identify whether or not many of the key components of the intervention were present. In addition to viewing the completed CSCs, the staff members who delivered the intervention completed an intervention fidelity checklist (see Appendix 14) following each intervention session. As no such checklist was already in existence, the researcher created this based on Gray's (1994b) guidance. The intervention diaries produced additional intervention fidelity information with regard to the level of exposure to the intervention that participants received.

3.10 Reliability and Validity

3.10.1 Internal Validity

Internal validity refers to the extent to which changes observed in a dependent variable can be ascribed to the effect of an independent variable, as opposed to other extraneous variables (Mertens, 2010). Table 3.8 displays the potential threats to internal validity within the present research along with the measures taken by the researcher to try to reduce these. Where threats were not able to be reduced, these are highlighted as limitations of the research in the discussion.

Certain SCEDs, notably A-B designs, have poor internal validity making inferring causal relationships particularly problematic (Barlow et al, 2009). It should be noted that SCEDs inherently attempt to reduce threats to internal validity through the administration of repeated measures over time (Horner et al, 2005). Kratochwill (1992) outlined how the internal validity of SCEDs can be improved by:

- basing the research on direct intervention
- basing the research on direct observational data
- collecting data from multiple outcome measures
- applying an intervention to several individuals who differ on a variety of characteristics
- demonstration of generalisation of an effect through the use of a multiple-baseline design.

These criteria were included in the design of this research.

Threat to Internal Validity	Design Consideration(s)
History: *Aspects that have changed in the environment other than the intervention*	• The participant inclusion criteria stated that participants must not already be receiving an intervention aimed specifically at the target behaviours of interest. Schools were asked to ensure that participants were not exposed to such interventions during the course of the research. • The researcher kept in regular contact with participating staff to ensure that up-to-date information about the participants' school environments was maintained.
Maturation: *Growth, change or development in participants unrelated to the intervention*	• The temporal sequencing element of the multiple-baseline design helped to rule out the likelihood that maturation could account for observed behaviour change. • The use of multiple cases can help to reduce this threat if intervention effects are demonstrated across different participants. • The research was conducted over a relatively short time period which perhaps limits the potential impact of maturation. However, the potential effects of maturation could not be completely controlled.
Instrumentation: *Changes in the instrument, observers, or scorers which may produce changes in outcomes*	• Observations were carried out by the same person throughout the research, in accordance with agreed schedules
Statistical regression: *Tendency for subsequent scores to regress towards the mean*	• This threat was reduced through repeated measures being taken over time to establish patterns of behaviour. • This may have remained a threat to the pre- and post- intervention questionnaire measures, but

	this data was only supplementary to the preliminary repeated measures data.
Experimental mortality: *Loss of participants*	• The researcher kept in regular contact with participating staff throughout the research to reduce the chances of drop-out.
Hawthorne effect: *Behaviour change in individuals due to participation in research, rather than manipulation of independent variables*	• Blinding procedures which guard against this threat to validity were not possible in this research for ethical reasons, therefore this remains a potential limitation.

Table 3.8: Threats to internal validity in this research and the measures taken by the researcher to reduce these (informed by Cook & Campbell, 1979).

3.10.2 External Validity

External validity refers to the extent to which research conclusions are more generally applicable beyond the specifics of the situation studied (Robson, 2011). There are obvious limitations in establishing wide generality from the results of SCEDs which focus on individuals (Barlow et al, 2009). However, SCEDs provide a suitable experimental method for research involving heterogeneous populations, such as those with ASC. The external validity of SCEDs can be enhanced through systematic replication (Horner et al, 2005) and this research involved multiple participants which went some way towards achieving this. Horner et al (2005) also highlight how providing details about the sample and context can help indicate who the intervention may be most effective for.

Whilst it was important to acknowledge that this research would not be able to make a persuasive case for the gerneralisability of the outcomes observed, the researcher considered that there remained worth in an in-depth investigation of the impact of CSCs on individual pupils as this would add to the limited

existing body of research, even if conclusions would necessarily be exploratory in nature.

3.10.3 Reliability

Reliability refers to the stability or consistency with which something is measured (Robson, 2011). In this research, the main measurement device was structured observation, with supplementary measures being taken via a self-completion pre-and post-intervention questionnaire. There exists a range of potential threats to reliability which could have impacted on these measures, as discussed below.

3.10.3.1 Reliability of Structured Observations

Robson (2011) outlines a number of threats specific to the reliability of observational measures. These are displayed in Table 3.9 along with the measures taken by the researcher to attempt to reduce the threats.

3.10.3.2 Inter-Observer Agreement

Calculating inter-observer agreement (IOA) (the extent to which two or more observers obtain the same results when measuring the same behaviour at the same time) is a widely used method of establishing the reliability of observational measures, although it remains possible that the observer themselves can be affected by testing for IOA (Robson, 2011). Although there is no set standard, Friman (2009) suggested that it is conventional for IOA to be assessed for at least 15% of observation sessions. A slightly higher figure has been proposed by Kratochwill et al (2013) who suggest that IOA should be collected for 20% of the data within each phase.

For four participants, a selection of observations were conducted jointly by a staff member and the researcher in order to assess IOA levels. Due to logistical and time constraints, the researcher experienced difficulties in achieving the higher figure of 20% proposed by Kratochwill et al (2013). Therefore, the

researcher aimed to achieve a minimum of 15% joint observation sessions and this was achieved in most cases. Further information is available in section 4.8.

Potential Threat to Reliability	Measure(s) Taken to Reduce Threat
Reactivity: *The extent to which an observer affects the situation under observation.*	This was deemed to be an issue when the researcher themselves was observing as their presence would not have been typical within the classroom. In these situations, the researcher aimed to be as unobtrusive as possible, keeping at a distance. Over time, it was expected that participants would become accustomed to the presence of the researcher and therefore less likely to behave differently to how they typically would.
Observer drift: *Changes in the way an observer uses an observation schedule, e.g. if increased familiarity with its use makes it easier to 'see' examples of categories.*	Operational definitions of behaviours were agreed upon at the outset to maintain consistency in recording. Joint observations were conducted to assess reliability (more information regarding this is supplied in section 3.10.3.2).
Expectancy effects: *Observers coding behaviour after an intervention may 'expect' to see changes.*	Use of operational definitions of behaviour aimed to ensure that only the target behaviour was captured consistently. This threat could not be fully removed as blinding procedures were not used.

Table 3.9: Threats to the reliability of the observational measures used and the measures taken by the researcher to reduce these.

3.10.3.3 Reliability of Pre- and Post-Intervention Measures

Single point data considered alone has numerous threats to validity and reliability (Robson, 2011). However, this data was supplementary to that obtained through the repeated measures and was included for data triangulation

purposes. It was considered that such information may be useful in cases where the outcomes of the repeated measures were unclear.

The key reliability issue relating to the use of the pre- and post-intervention questionnaire is respondent bias (Robson, 2011). Respondent bias is a general term for a range of cognitive biases which may affect the way in which a respondent answers a question. As the staff members who completed the questionnaires were aware of the purpose of the research, it was possible that they may have answered in a way which they thought would please the researcher. Although this is a risk that could not be completely controlled for, all staff members were encouraged to provide their honest opinions.

3.11 Ethical Considerations

Throughout the planning and implementation of this research, the researcher referred to The British Psychological Society Code of Human Research Ethics (BPS, 2010) and the University of Nottingham Code of Research Conduct and Research Ethics (2013). Approval from the University of Nottingham Ethics Committee was obtained on 31st March 2015 (see Appendix 15). The key ethical considerations related to this research are summarised in Table 3.10.

Ethical Issue	Design Considerations
Informed consent	• Informed consent was gained from all involved (participants, parents and staff). In addition to providing written information detailing the nature of the research, the researcher met directly with school staff to provide an opportunity for further discussion and to ask any questions that they had about the research. • Consent was gained from participants (as deemed appropriate to their developmental level) after a familiar member of school staff explained the research to them in an age- appropriate manner. It was explained to the participants that they could ask any questions that they wished about the research. • The researcher provided their contact details should any party wish to seek further information at any point throughout the research.
Right to withdraw	• It was made clear during the consent gaining stage that all involved had a right to withdraw from the research at any stage without giving a reason.
Confidentiality	• All the data collected was kept anonymous (e.g. through using participant numbers rather than names) and confidential. The researcher assured that no child or school would be identifiable in the research write-up.
Risk	• It was considered that there may be the potential for participants to become distressed when talking about a situation that may have been difficult for them. It was therefore agreed with school staff that any concerns regarding the well-being of participants that was noted during the course if the research would be raised immediately.
Debriefing	• A debrief statement was provided to participants, parents and participating school staff as soon as data gathering was completed. • It was planned for stakeholders to be provided with a written summary of the results of the research

Honesty and Integrity	• In order to capture an authentic a picture as possible of the frequency of the target behaviours of interest, the observations that were conducted were not made overt to the participants. However, participants were informed during the consent gaining stage that information about their behaviour in school was going to be collected • Participants were, through a process of debriefing, provided with further information about the nature of the observations that were conducted over the course of the research period.
Good communication	• The researcher liaised regularly with school staff throughout the research period. Staff and parents were provided with the researcher's contact details should they wish to make contact at any point. • All involved were also informed that, following the completion of the research, they would receive a summary of the findings.

Table 3.10: Ethical considerations for this research

Chapter 4 - Results

4.1 Introduction to Results Chapter

This chapter begins with a discussion of different methods of analysing SCED data, including their relative strengths and limitations. The analysis approach used in this research, and the rationale for this, is then outlined.

The results of the present study are then presented and analysed, with the results for each participant being considered individually. Data regarding the reliability of the visual analysis and the observational measures is also presented alongside intervention fidelity data.

4.2 Visual Analysis

Traditionally, single-case researchers have relied on visual analysis of data to determine whether or not there is evidence of a causal relationship between an independent variable and a dependent variable (Kratochwill et al, 2013). Essentially, this involves displaying a participant's performance in the different phases in graphical form and then comparing the phases (Robson, 2011). A stable baseline makes analysis much easier (Barlow et al, 2009; Kazdin, 2003).

Kratochwill et al (2013) outline a set of standards for conducting visual analysis. Their process begins with the following four steps:

1.   documentation of a predictable and stable baseline
2.   examination of data within each phase to assess within phase patterns
3.   comparison of data between adjacent phases to assess whether the independent variable can be tied to an 'effect'
4.   integration of information from all phases to determine if there are at least three demonstrations of an effect at different points in time (N.B. in this research, there were only two opportunities to assess intervention effects within each case).

To assess if an intervention effect is likely present, there are six features of the data that can be examined (Kratochwill et al, 2013), as outlined in *Table 4.1*.

Feature	Definition of the Feature
Level	The overall average (mean) of the measures within a phase

In the present research, a <u>decrease</u> in mean level was indicative of a desirable intervention effect. |
| Trend | The slope of the best-fitting straight line for the measures within a phase |
| Variability | The range, variance or standard deviation of the measures about the best-fitting line.

Low levels of variability are indicative of a more stable and reliable data set. |
| Immediacy of effect | The change in level between the last three data points in one phase and the first three data points of the next

The more rapid the effect, the more convincing the inference that change can be attributed to the introduction of the independent variable. |
| Overlap | The proportion of data from one phase that overlaps with data from the previous phase.

The smaller the proportion of overlap, the more convincing the demonstration of an effect. |
| Consistency of data patterns across similar phases | Examining the extent to which there is consistency in the data patterns from all phases within a condition (i.e. all baseline phases / all intervention phases).

Greater consistency provides greater confidence in concluding an intervention effect. |

Table 4.1: Features of visual analysis of SCED data informed by Kratochwill et al (2013)

Some of these criteria (e.g. trend) can be used to make predictions about expected patterns of behaviour had an intervention not been introduced. This can then be compared to the actual data collected (Kratochwill et al, 2013). It is the role of the data analyst to judge the extent to which changes in the criteria are present across the phases (Kazdin, 2003).

4.2.1 Reliability of Visual Analysis

Whilst visual analysis is thought to sensitise oneself to the critical properties of the data collected, the reliability of it has been questioned due to its subjective nature and common disagreement between raters (Kazdin, 2003; Brossart, Parker, Olson & Mahadevan, 2006). Visual analysis may lack sensitivity, leading to intervention effects only being identified when they are particularly marked, perhaps missing meaningful (albeit more modest) effects, with this increasing the likelihood of type II errors (Kazdin, 2003; Parker & Hagan-Burke & Vannest, 2007).

Data sets containing unstable baselines are much harder to interpret through visual analysis (Brossart et al, 2006), as are data sets which display a baseline trend similar to that which is predicted to follow the introduction of an intervention (Brossart, Vannest, Davis & Patience, 2014; Kazdin, 2003).

A further limitation of visual analysis relates to the issue of autocorrelation (or 'serial dependence') within time series data (Barlow et al, 2009). Autocorrelation refers to the correlation between data points separated by different time intervals (Barlow et al, 2009). Whilst there are conflicting findings about the extent to which autocorrelation is present in behavioural data, Perrin (1998) cautions visual analysts that what may appear to be a clear pattern of change across phases may actually have been quite predictable due to inherent correlation. Failing to take account of autocorrelation increases the likelihood of a type I error being made (Barlow et al, 2009).

The reliability of visual analysis can be assessed and enhanced through calculating a level of agreement between two independent visual analysts (Brossart et al, 2006). Inter-rater reliability was assessed in this research and further information with regards to this can be found in section 4.9.

4.3 Statistical Analysis

Statistical analysis has been proposed as a useful supplement to visual analysis, especially in cases where there is an unstable baseline and general variability within data sets (Kazdin, 2003). The inclusion of statistical analysis of SCED

data is believed to be increasing due to the drive for more objective and statistically significant outcomes within intervention outcome research (Parker et al, 2007). Whilst the use of statistical methods allows for enhanced sensitivity to detect intervention effects (as opposed to visual analysis alone) it is important to remember that they are more likely to capitalise on chance and therefore result in a type I error (Nock et al, 2007; Todman & Dugard, 2001).

Inferential statistical tests commonly used in group designs, such as the t and F tests, are not appropriate for single case studies due to the possibility of autocorrelation, which would violate the independence of error assumption of such tests (Nock et al, 2007). Additionally, as SCEDs are characterised by small sample sizes, they are unlikely to meet the parametric assumptions of conventional statistical tests of difference. There are several non-parametric alternatives which can be used to analyse single case data SCEDs, many of which are dependent upon certain characteristics of the data set. A number of these were considered in order to determine whether or not they were an appropriate means of data analysis in this research.

Interrupted Time-Series Analysis (ITSA):

The use of time-series analysis has become increasingly popular over time and is able to control for the autocorrelation of data points, before using a t test to assess change (Barlow et al, 2009). For ITSA to be a viable option, the data needs to meet a range of assumptions including that the series is equally spaced, consists of at least 50 observations and has no missing values. As these assumptions are not met by the data in the present research, this was not a suitable analysis method to use.

Randomisation Tests:

Randomisation tests make no assumptions about the distribution of the data. However, crucial to the use of such tests is the concept of randomisation. In SCEDs, this concept translates to the random assignment of the independent variable to measurement occasions (Onghena & Edgington, 2005). It is possible that an intervention effect can intermingle with naturally occurring extraneous factors to create an effect on the dependent variable, therefore

obscuring the 'true' value of the intervention effect. Hence, randomisation is advantageous as it allows for a more reliable estimate of the effect of the intervention (Barlow et al, 2009).

However, a number of issues may adversely affect the statistical power of randomisation tests including a limited number of ways that the intervention can be randomised, autocorrelation of data and a delayed response and/or intervention carryover effects (Onghena & Edgington, 2005). Computational burdens and a lack of suitable software is a further consideration (Barlow et al, 2009). For these reasons, and as a result of the non-random assignment of the intervention to measurement occasions, randomisation tests were not appropriate in this research.

4.3.1 Effect Size

The use of inference testing alone is not deemed sufficient for certain conclusions in SCEDs, as a calculation of the 'significance' of findings does not necessarily ensure that results are actually meaningful at the level of clinical relevance (Barlow et al, 2009). This argument has led to the development of a range of effect size measures.

Effect size is an index of the strength of association between intervention and outcome and provides a measure of practical, rather than statistical, significance (as statistical significance conflates effect size and sample size) (Brossart et al, 2006). It is said to be a beginning point for overlaying social value judgements by teachers and EPs (Parker, Vannest & Brown, 2009), a particularly pertinent consideration given the context of the present research. Supplementing visual analysis with an effect size can provide standardised and reliable results that contribute to evidence-based practices (Vannest & Ninci, 2015) through providing improved measurement precision when results are not large and obvious, as well as providing an objective summary when visual judgments do not agree (Parker et al, 2007). It is argued that single-case researchers should always conduct both visual analysis and effect size analysis, as they should reinforce and inform each other (Brossart et al, 2014).

Several methods are available for measuring effect size in SCEDs which broadly fall into three categories: regression models (e.g. Cohen's r and R^2), standardised mean difference (e.g. Cohen's d and Hedge's g) and non-overlap indices (Ross & Begeny, 2014). Manolov & Solanas (2008) found that autocorrelation least affected effect sizes calculated by percentage of non-overlapping data indices. Non-overlap methods are straightforward, distribution-free and non-parametric and have the benefit of being visually accessible and blending well with visual analysis (Parker, Vannest, Davis & Sauber, 2011). An outline of a number of non-overlap effect size indices is provided below.

4.3.1.1 Percentage of Non-Overlapping Data (PND)

Calculating PND (Scruggs, Mastropieri & Casto (1987) involves identifying the most extreme score in the baseline data series then identifying the number of data points within the intervention phase which exceed this. PND is one of the most widely used and straightforward of the quantitative methods. However, it is thrown off by outlier scores and cannot address trend. Additionally, it cannot be used for significance testing or to calculate confidence intervals, both of which it is widely suggested by researchers should be reported if possible (Vannest & Ninci, 2015). Several researchers recommend that PND is not used (Kratochwill et al, 2013, Parker et al, 2009).

4.3.1.2 Percentage of Data Exceeding the Median (PEM)

PEM (Ma, 2006) relies on finding the median score for the baseline data and then identifying the number of data points in the intervention phase which overlap with this. Whilst it addresses more of the baseline data than PND does, PEM is affected by high variability as, in such cases, the median will not be an adequate summary of the score distribution and results will therefore be distorted (Parker et al, 2007). As PEM does not consider trend, it is not optimal to use this method when trend is present in the baseline (Vannest & Ninci, 2015). Additionally, PEM lacks sensitivity to the magnitude of an intervention effect, as both slightly non-overlapping and largely non-overlapping data sets would produce the same 100% effect (Ma, 2006; Vannest & Ninci, 2015).

4.3.1.3 Improvement Rate Difference (IRD)

Well established within medical research, where it is known as the 'risk reduction technique', IRD (Parker et al, 2009) is defined as the improvement rate of the intervention phase minus the improvement rate of the baseline phase, therefore the effect size demonstrates a proportion of improvement rather than a percentage of change (Vannest & Ninci, 2015). IRD tends to be more robust than PND or PEM as it takes account of more data in the calculation, but like the others this method cannot address trend (Vannest & Ninci, 2015). The outcomes of an IRD analysis also appear to be somewhat dependent on the length of a data series (Manolov, Solanas, Sierra & Evans, 2011).

4.3.1.4 Tau-U

Tau-U (Parker et al, 2011) is derived from Kendall's Rank Correlation and the Mann-Whitney between groups U-test. It involves a pairwise comparison between each data point in the baseline phase and each data point in the intervention phase to identify if the earlier data point is larger, smaller or equal to the latter. Tau-U can be conceptualised as the percentage of non-overlap between phases or the percentage of data showing improvement between phases (Parker et al, 2011). Tau-U uses every data point in each phase in the analysis, is only somewhat influenced by autocorrelation and has the ability to adjust for trend, thus addressing the issues that are problematic for the other effect size measures outlined above. It is also able to handle smaller data sets and has more power than other non-overlap methods (Brossart et al, 2014). For these reasons, it has been proposed that Tau-U is the better performing non-parametric method for analysing single-case data presently available (Brossart et al, 2014).

4.3.1.5 Limitations of Effect Sizes

Effect sizes will vary depending on the analytic method used, meaning interpreting the significance of outcomes obtained can prove difficult (Cohen, 1992). As there is presently no agreed upon criteria for statistical analysis of

SCED data (Kratochwill et al, 2013), decisions regarding whether to use visual analysis, statistical analysis, or both should be made based on the design of the study and the researcher's goals (Nock et al, 2007). If an effect size is to be calculated, it is important to choose a method that makes the most sense with regards to the characteristics of the data, which in SCEDs often includes small data sets, variability, small or gradual behaviour change, and trend (Brossart et al, 2014). It is also important to recognise that whilst statistical methods are able to produce effect sizes, they cannot factor in the multiple ways that context can impact the interpretation of one's data (Brossart et al, 2014). Numbers alone, without a contextualised interpretation, lead to an incomplete picture for evaluating the practical strength of an intervention (Vannest & Ninci, 2015).

Various authors (e.g. Brossart et al, 2014; Parker et al, 2007) strongly recommend that confidence intervals around the effect size are reported in order to assess its reliability. If a data set is short, the effect size will likely have low reliability or dependability, regardless of its possibly large and visually convincing size. This will be apparent if the effect size is bracketed by a wide confidence interval (Parker et al, 2007).

4.4 Data Analysis Used in This Research

4.4.1 Analysis of Repeated Measures
4.4.1.1 Visual Analysis

The effectiveness of the CSC intervention was initially assessed using visual analysis. Graphs were created using Microsoft Excel following guidance from Dixon et al (2009). The researcher considered a range of change factors in accordance with the criteria recommended by Kratochwill et al (2013) (outlined in Table 4.1). It should be noted that 'consistency of patterns across similar phases' was not considered as there was only one baseline and one intervention phase per behaviour.

On the graphs, dotted lines were used to indicate where there was missing data. This included holiday periods as well as missing data related to factors such as

participant or staff absence. The dotted lines were included to aid the researcher in their visual exploration of overall data patterns. Alongside the graphs, quantitative and qualitative descriptions of the change factors are provided.

The reliability of the visual analysis was also assessed (further details can be found in section 4.9).

4.4.1.2 Effect Size Analysis

In addition to visual analysis, effect sizes were calculated using the Tau-U non-overlap index. As stated previously, inclusion of statistical analysis is now deemed desirable in SCEDs in order to provide a more objective measure of intervention impact, especially in cases where undesirable data characteristics are present. It was clear that baseline trend and data variability were present in a number of the data sets in this research. This provided sufficient reason to carry out additional analyses in order to have more confidence in the reliability of the results. Tau-U was deemed a highly appropriate method as it would accommodate for baseline trend.

In addition to the reason outlined above, in order to adequately address Research Question 2 it was considered that an objective measure would be important when exploring differential outcomes in relation to the frequency of intervention implementation. It was considered that visual analysis alone may not have been sufficient to identify differential effects, especially in the presence of variable data sets or smaller changes.

In addition to an effect size, confidence intervals and p-values can be calculated by the Tau-U method and therefore these are also reported in order to provide an estimate of the probability of chance occurrence of the effect size and an estimate of the error based on the confidence intervals (Brossart et al, 2014). Parker et al (2011) suggest that inference testing is important with short data series and there were a number of these within the present research. The significance threshold was set at $p < 0.05$. Tau-U analysis was conducted using

the free software available to researchers at www.singlecaseresearch.org. The Tau-U output tables for each participant can be viewed in Appendix 16.

It should be noted that there is ongoing exploration into the most appropriate effect size benchmarks to use with the growing range of single-case effect size measures, making it difficult to establish a clear criteria. In this research, the effect sizes were interpreted in line with benchmarks proposed originally by Cohen (1988) and more recently deemed appropriate for use with the Tau statistic by Ferguson (2009): ≥ 0.20 was interpreted as a small effect; ≥ 0.50 a moderate effect; and ≥ 0.80 a large effect. These benchmarks appeared the most appropriate available at the present time and provided a means of comparing results across participants. However, the researcher acknowledged the limitations of attributing too much weight to terms such as 'small', 'moderate' and 'large' without due consideration of contextual factors which contribute towards evaluations of the practical strength of any intervention (Vannest & Ninci, 2015).

4.4.2 Analysis of Pre- and Post-Intervention Measure

Quantitative data from the pre- and post-intervention questionnaires is presented in a table and the difference in ratings calculated to highlight any changes. The rating given for the perceived effectiveness of the CSC intervention is considered alongside the results of the visual analysis and Tau-U analysis. Where additional comments were noted, these are provided as additional contextual information and will be further referenced where appropriate in Chapter 5.

4.5   Results of Each Case Study

The results of each case study are presented individually with the results of the visual analysis, Tau-U effect size analysis and the pre- and post- intervention data all being considered. Where appropriate, numbers have been rounded to two decimal places.

4.5.1   Case Study One: Jack

4.5.1.1   Pupil Profile

Age at start of data collection: 6 years and 10 months

TB1: Not waiting his turn to speak

TB2: Touching/squeezing peers

Order of Conditions: Jack started with the CSCW intervention (targeting TB1) followed by CSCS (targeting TB2).

Numbers of CSCs completed in CSCW condition: 3

4.5.1.2 Outcomes of Repeated Measures
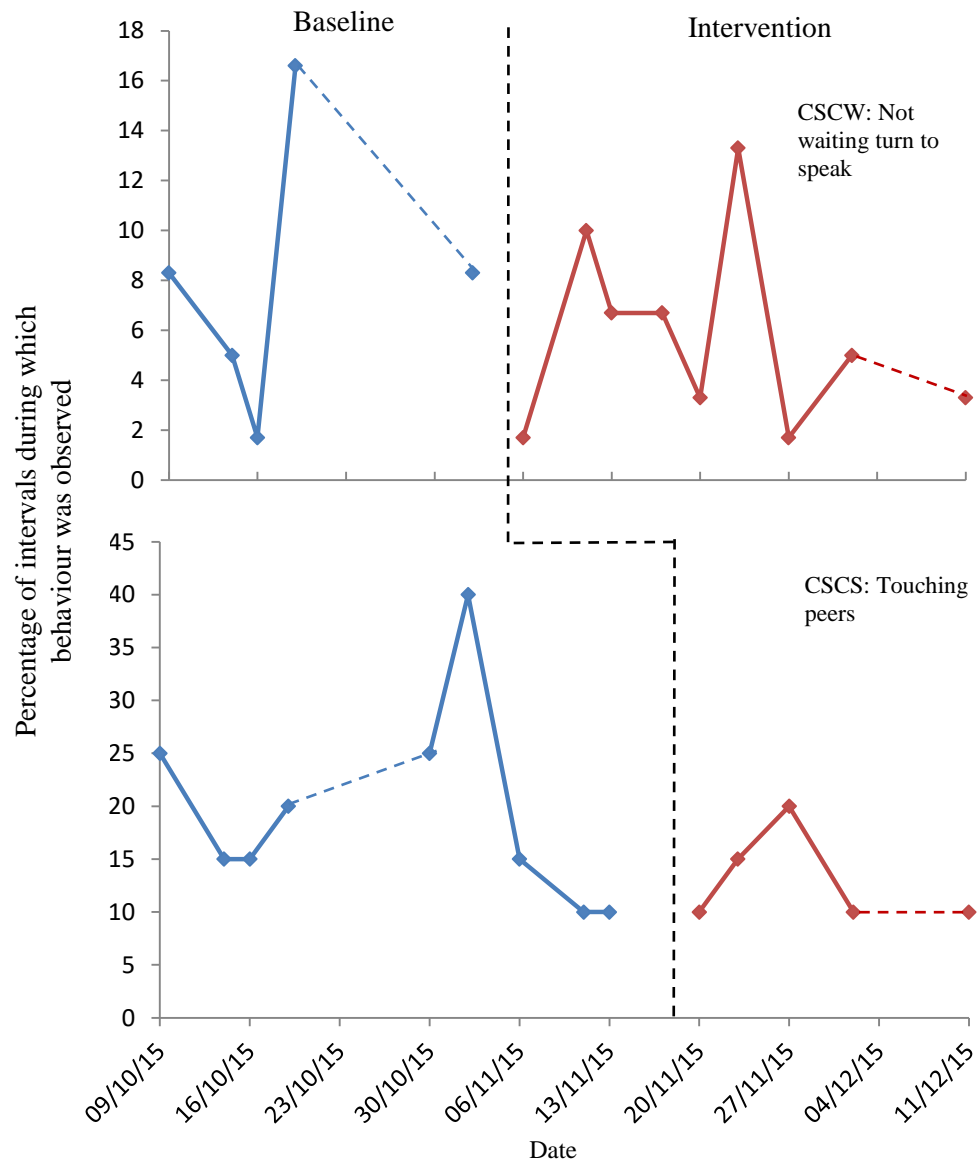
Graph 4.1: A scattergraph showing the frequency of Jack's target behaviours across baseline and intervention phases
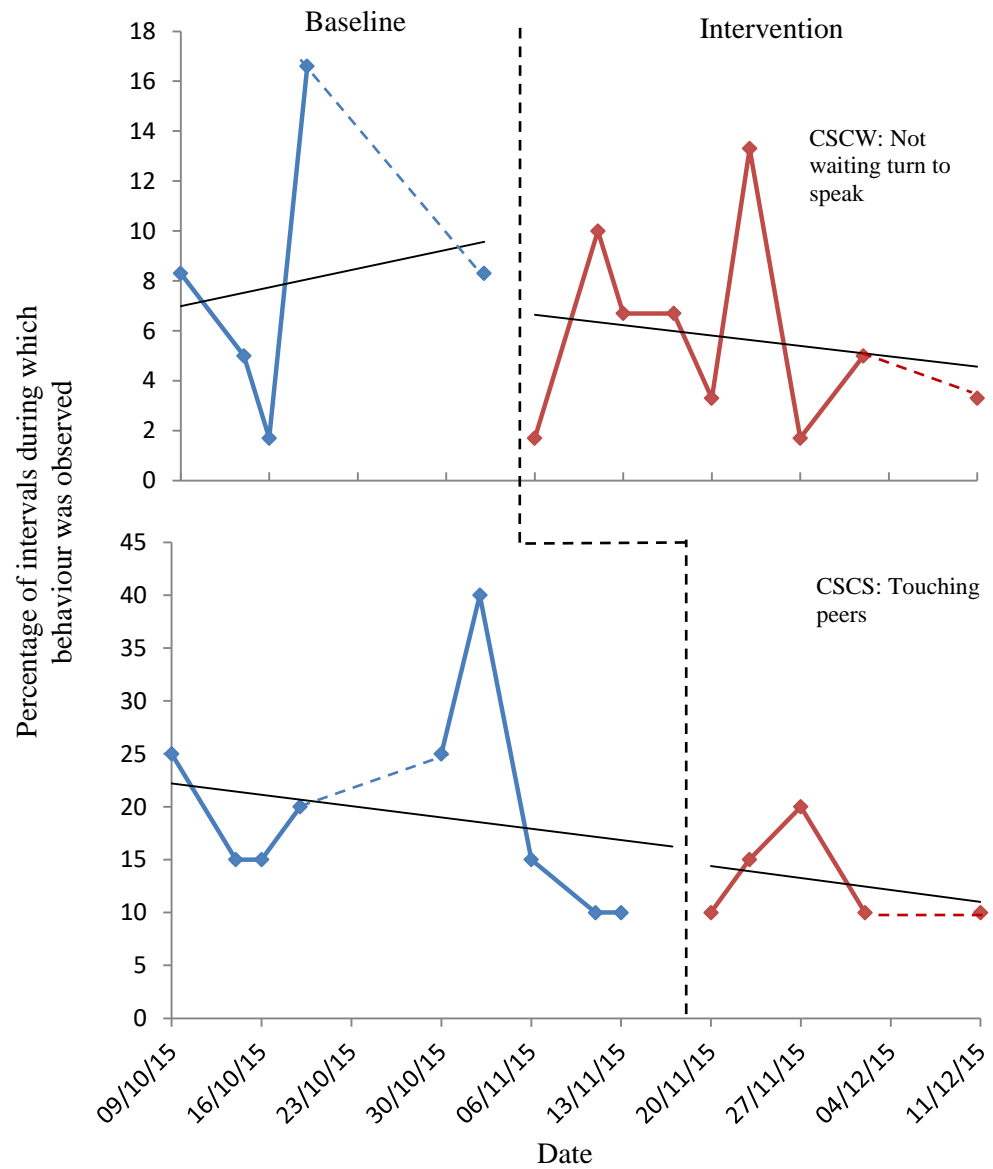
Graph 4.2: A scattergraph showing the frequency of Jack's target behaviours across baseline and intervention phases with trendlines

Graph 4.3: A scattergraph showing the frequency of Jack's target behaviours across baseline and intervention phases with mean lines and variability lines

Characteristic	Description of outcome for TB1 (CSCW)	Description of outcome for TB2 (CSCS)
Level	Mean of baseline phase = 7.98% Mean of intervention phase = 5.74% There was a mean level decrease of 2.24% between baseline and intervention.	Mean of baseline phase = 19.44% Mean of intervention phase = 13% There was a mean level decrease of 16.44% between baseline and intervention.
Trend	Graph 4.2 shows a slight upward trend in the baseline phase and a slight downward trend in the intervention phase.	Graph 4.2 shows a slight downward trend in the baseline phase and this trend continues in the intervention phase.
Variability	Graph 4.3 shows a fairly high level of variability of scores in the baseline phase (range = 14.9 ; S.D^3 = 5.54). There is slightly less variability in the intervention phase (range = 11.6; S.D = 3.90).	Graph 4.3 shows high variability in the baseline phase (range = 30; S.D = 9.50). There is much less variability in the intervention phase (range = 10; S.D = 4.47)
Immediacy of Effect	There is an immediate decrease from the final data point of the baseline phase and the first data point of the intervention phase but this is not maintained.	There is an overall (undesirable) increase level between the last three data points in the baseline phase and the first three in the intervention phase.
Overlap	All data points (100%) in the intervention phase overlapped with those in the baseline phase.	All data points (100%) in the intervention phase overlapped with those in the baseline phase.

Table 4.2: Summary of the outcome of visual analysis for Jack's frequency of behaviours graphs

---

[3] Standard deviation

4.5.1.3 Results of Tau-U Analysis

Target Behaviour	Tau-U Effect Size	Confidence Interval (90%)	P-Value
1 (CSCW)	-0.29	-0.84 < > 0.26	0.39
2 (CSCS)	-0.22	-0.77 < > 0.33	0.51

Table 4.3: Results of the Tau-U effect size analysis with confidence intervals and p-values (Jack)

4.5.1.4 Summary of Outcomes of Repeated Measures

TB1: CSCW

The visual analysis suggested a minimal intervention effect. The upward trend in the baseline phase is followed by a desirable downward trend during the intervention phase and there is a small decrease in the mean (-2.24%) between phases. However, variability within both phases and 100% data overlap make it difficult to draw firm conclusions.

The Tau-U analysis produced an effect size of -0.29, which equates to a 'small' effect in the desired direction (i.e. a reduction in the frequency of behaviour). This was a non-significant finding (p = 0.39).

TB2: CSCS

The visual analysis suggested a possible change in relation to the frequency of TB2. There was a desirable mean level decrease (-16.44%) between baseline and intervention and much less variability in the intervention phase. However, an unstable baseline makes analysis more difficult and there is a downward trend in the baseline phase suggesting that some improvement could be predicted regardless of the intervention. In addition, there is 100% overlap of data.

The Tau-U analysis produced a small effect size of -0.22. This was a non-significant finding (p = 0.51).

In considering the multiple-baseline graph as a whole, this does not reliably demonstrate that a change occurred only when the intervention was directed at the target behaviour, as there was a concurrent decrease in the frequency of TB2 at the time of intervention implementation for TB1, although the generally unstable baselines make identifying clear changes very difficult.

4.5.1.5   Results of the Pre- and Post- Measure

TB1: CSCW

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	8	7
Post-intervention	7	7
Difference	-1	0

Table 4.4: Staff member ratings given on the pre- and post-intervention questionnaires for TB1 (Jack)

Table 4.4 shows that there was little or no change in the staff member's ratings of Jack's TB1 in relation to how challenging/disruptive they find it to be and how much of an impact it has on his learning and/or social relationships. This data is somewhat supportive of the conclusions of the visual analysis and effect size analysis, which suggested only a small possible intervention effect.

In contrast to their behaviour ratings, the staff member rated the CSCW intervention at '6' for effectiveness in addressing TB1 and commented that:

"The CSC was useful when bringing the target behaviour to the attention of the child as sometimes he didn't realise he was doing it" but that *"after a few weeks of targeting the same target behaviour, he seemed to get a little bored with the process".*

TB2: CSCS

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	7	7
Post-intervention	6	7
Difference	-1	0

Table 4.5: Staff member ratings given on the pre- and post-intervention questionnaires for TB2 (Jack)

There was little or no change in the staff member's ratings of the impact of Jack's TB2 in relation to how challenging/disruptive they find it to be and how much of an impact it has on his learning and/or social relationships. This data is somewhat supportive of the conclusions of the visual analysis and effect size analysis, which suggested only a small possible intervention effect.

In contrast to their behaviour ratings, the staff member rated the CSCW intervention at '6' for effectiveness in addressing TB1 and commented that:

"Using the CSC was good for putting the participant in the shoes of the children involved and to talk about how they might be feeling".

4.5.2 Case Study Two: Daniel

4.5.2.1 Pupil Profile

Age at start of data collection: 6 years and 6 months

TB1: Asking time-related questions

TB2: Waving finger in people's faces

Order of conditions: Daniel started with the CSCS intervention (targeting TB1) followed by the CSCW intervention (targeting TB2).

Number of CSCs completed in CSCW condition: 2 (intervention not delivered in accordance with the proposed schedule).

4.5.2.2 Outcomes of Repeated Measures
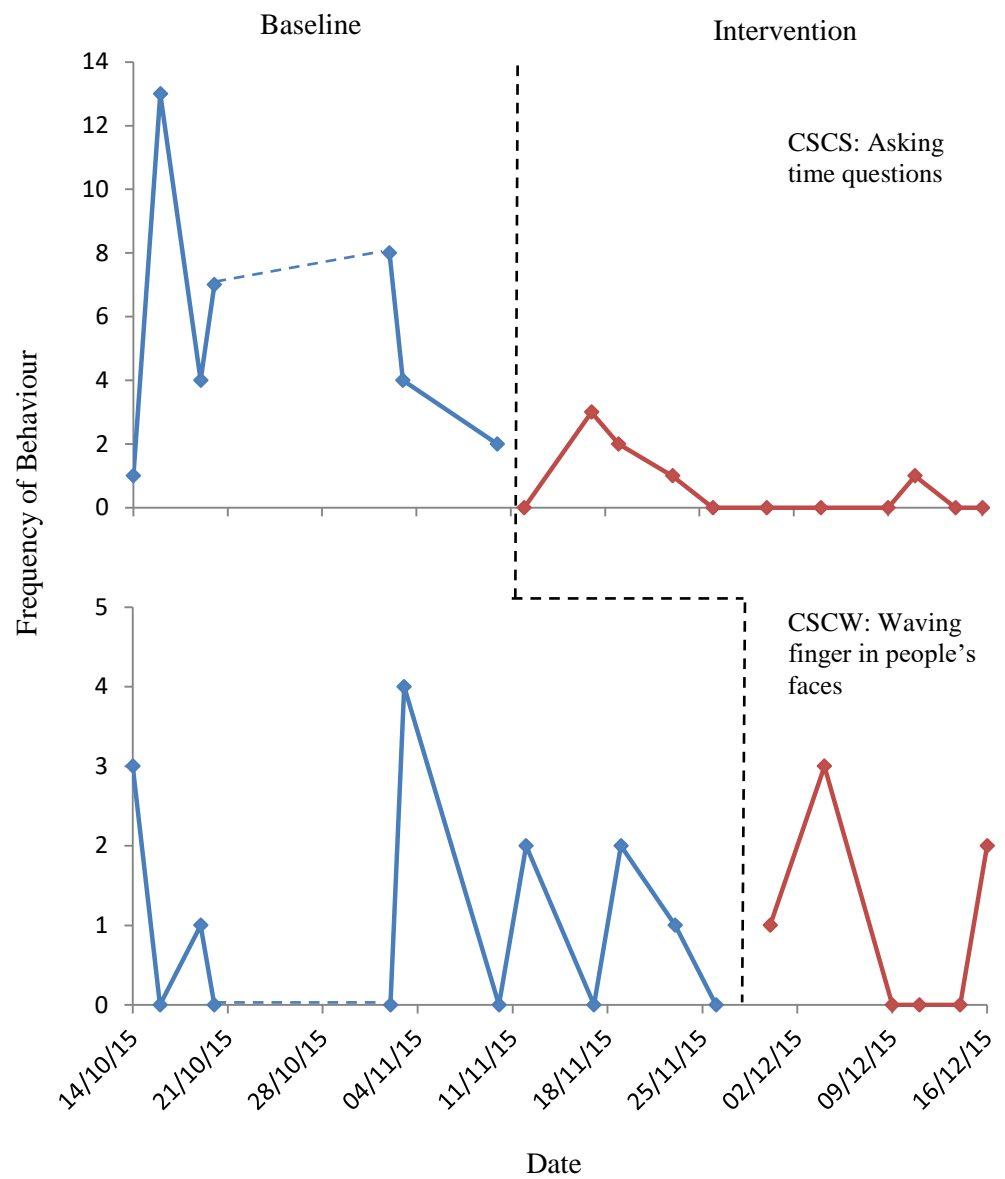
Graph 4.4: A scattergraph showing the frequency of Daniel's target behaviours across baseline and intervention phases
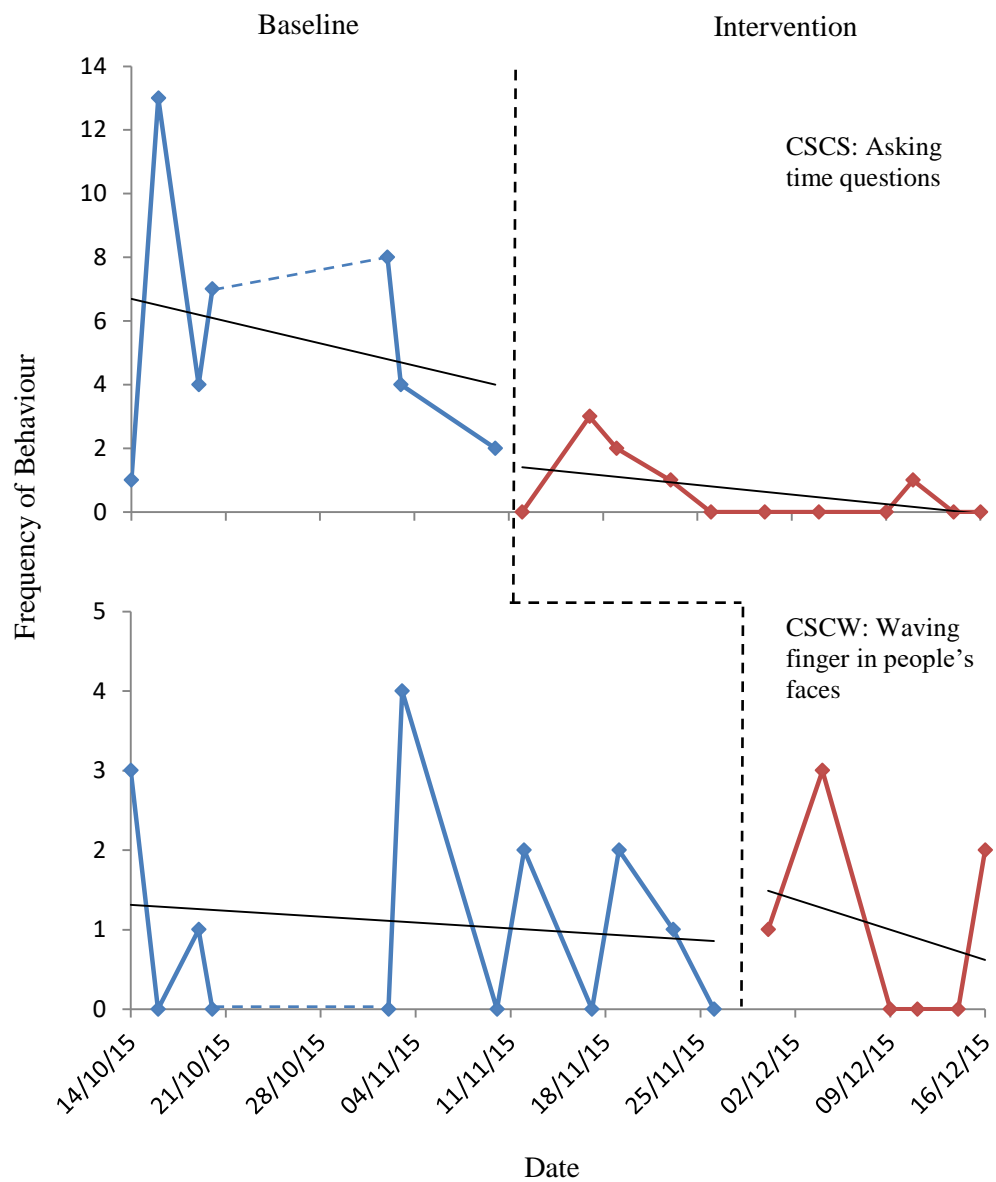
Graph 4.5: A scattergraph showing the frequency of Daniel's target behaviours across baseline and intervention phases with trendlines

Graph 4.6: A scattergraph showing the frequency of Daniel's target behaviours across baseline and intervention phases with mean lines and variability lines

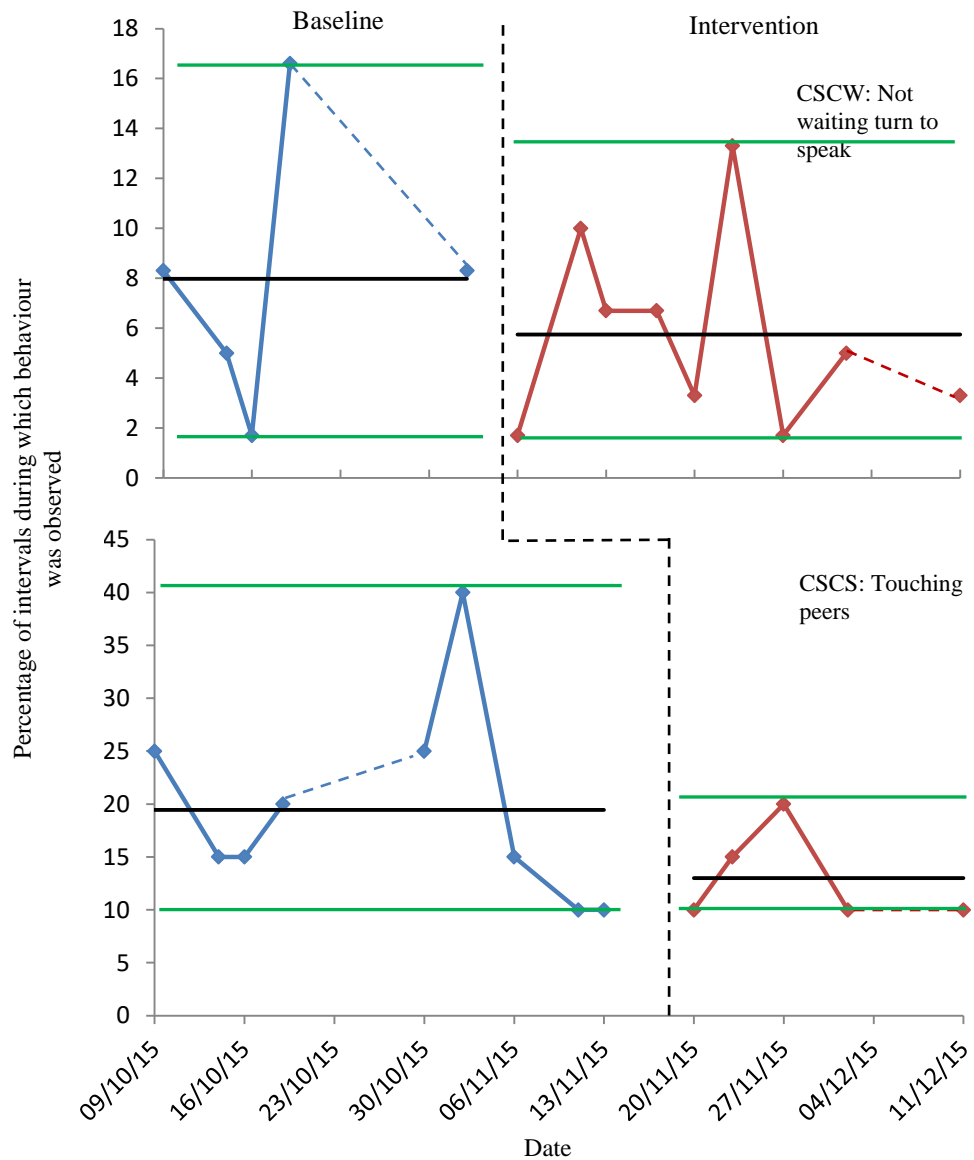| Characteristic	Description of outcome for TB1 (CSCS)	Description of outcome for TB2 (CSCW)
Level	Mean of baseline phase = 5.57 Mean of intervention phase = 0.64 There was a mean level decrease of 4.93 between baseline and intervention.	Mean of baseline phase = 1.08 Mean of intervention phase = 1 There was a mean level decrease of 0.08 between baseline and intervention.
Trend	Graph 4.5 shows a downward trend in the baseline phase and a similar downward trend in the intervention phase.	Graph 4.5 shows a very slight downward trend in the baseline phase and a more pronounced downward trend in the intervention phase.
Variability	Graph 4.6 shows a high level of variability in the baseline phase (range = 12; S.D = 4.12). There is much less variability in the intervention phase (range = 3; S.D = 1.03).	Graph 4.6 shows some variability in the baseline phase (range = 4; S.D = 1.38). There is a similar level of variability in the intervention phase (range = 3; S.D = 1.26)
Immediacy of Effect	There is an immediate change in level following the final data point of the baseline phase, but this is not maintained over the subsequent two data points.	There is little difference in overall level between the last three data points in the baseline phase and the first three in the intervention phase.
Overlap	36.36% of the data points in the intervention phase overlapped with those in the baseline phase.	All data points (100%) in the intervention phase overlapped with those in the baseline phase.

Table 4.6: Summary of the outcome of visual analysis for Daniel's frequency of behaviours graphs

4.5.2.3   Results of Tau-U Analysis

Behaviour	Tau-U Effect Size	Confidence Interval (90%)	P-Value
1 (CSCS)	-0.88	-1.36 <> -0.41	< 0.01
2 (CSCW)	-0.01	-0.50 <> 0.47	0.96

Table 4.7: Results of the Tau-U effect size analysis with confidence intervals and p-values (Daniel)

4.5.2.4   Summary of Outcomes of Repeated Measures

TB1: CSCS

The visual analysis suggests an observable intervention effect. Although there is a moderate downward trend during the baseline phase, there is an immediate decrease in level at the very start of the intervention phase (although this is not maintained immediately after) and from here a downward trend continues, with the behaviour being almost eliminated by the end of the phase. There is a mean level decrease of almost 5 and significantly less variability in the intervention phase. The relatively low level of 36.36% overlap adds additional weight to the suggestion of a possible intervention effect. However, an unstable baseline makes analysis more difficult.

The Tau-U analysis produced an effect size of -0.88, which can be classed as a 'large' effect and this was a statistically significant result ($p < 0.01$).

TB2: CSCW

The visual analysis suggests there was no intervention effect. There was a negligible decrease in the mean (-0.08) and a similar level of variability between phases, plus 100% data overlap. The downward trend in the intervention phase is slightly steeper than that of the baseline phase although not sufficiently so to indicate any clear effect. This was confirmed by the Tau-U analysis which produced a negligible and non-significant ($p = 0.96$) effect size of -0.01.

It should be noted that many of the baseline scores for TB2 were already at or near the floor which proves troublesome for analysis as there was little room to show improvement. In addition, the CSCW schedule was not implemented as planned, with only two CSCs being completed within the intervention phase.

In considering the multiple-baseline graph as a whole, this does not reliably demonstrate that a change occurred only when the intervention was directed at the target behaviour, although the generally unstable baselines make identifying clear changes very difficult.

4.5.2.5 Results of Pre- and Post-Intervention Measures

TB1: CSCS

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	9	9
Post-intervention	1	1
Difference	-8	-8

Table 4.8: Staff member ratings given on the pre- and post-intervention questionnaires for TB1 (Daniel)

Table 4.8 shows that there was a substantial decrease (-8) in the staff member's ratings of Daniel's TB1 in relation to how challenging/disruptive they find it to be and how much of an impact it has on his learning and/or social relationships. This data is supportive of the conclusions of the visual analysis and effect size analysis.

The staff member rated the CSCS intervention at '9' for effectiveness in addressing TB1 and commented that Daniel:

"Took the speech bubbles very literally and had to be reminded that the good strategy to resolve the problem was the important part" but that *"overall, he responded in a very positive way to reach very satisfactory conclusions"*.

TB2: CSCW

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	6	7
Post-intervention	3	2
Difference	-3	-5

Table 4.9: Staff member ratings given on the pre- and post-intervention questionnaires for TB2 (Daniel)

Table 4.9 shows that there was a modest decrease (-3) in the staff member's ratings of Daniel's TB2 in relation to how challenging/disruptive they find it to be, and a slightly larger decrease (-5) in relation to how much of an impact it has on his learning and/or social relationships. This data is not a very good fit with the conclusions of the visual analysis and effect size analysis, which suggested no real behaviour change.

In contrast to the effect size obtained, the staff member rated the CSCS intervention at '6' for effectiveness in addressing TB1 and commented that Daniel:

"Was open to illustrating the problem and keen to convert behaviours to please adults and children around him".

4.5.3   Case Study Three: Robert

4.5.3.1   Pupil Profile

Age at start of data collection: 8 years and 4 months

TB1: Not starting tasks independently

Order of conditions: Robert was exposed to the CSCS intervention only

4.5.3.2 Outcomes of Repeated Measures
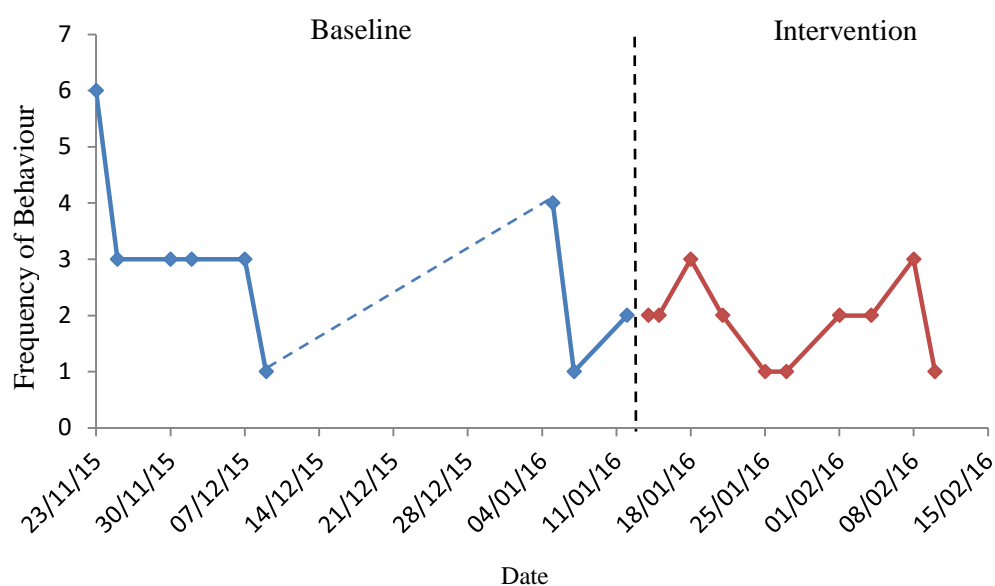
Graph 4.7: A scattergraph showing the frequency of Robert's target behaviours across baseline and intervention phases
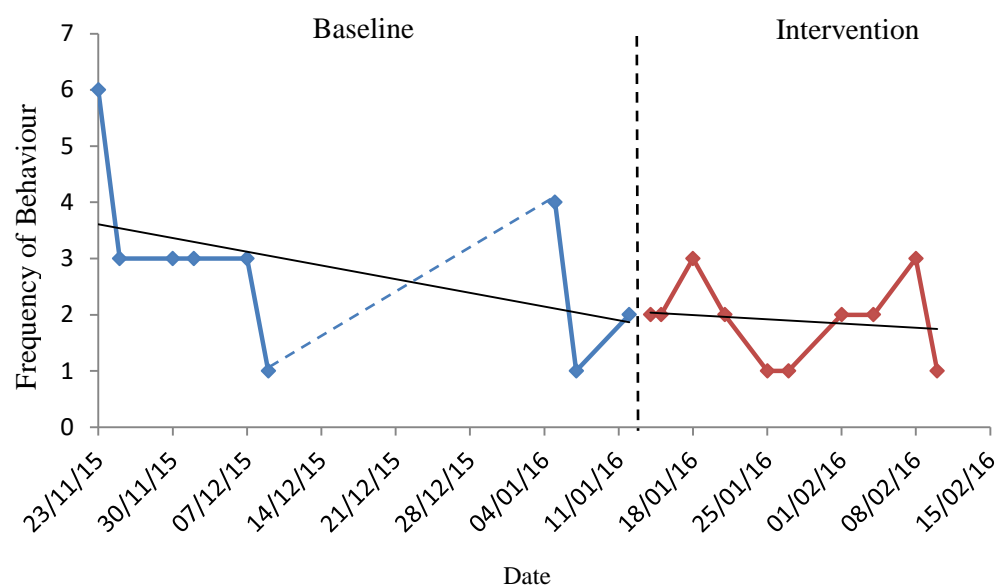
Graph 4.8: A scattergraph showing the frequency of Robert's target behaviours across baseline and intervention phases with trendlines

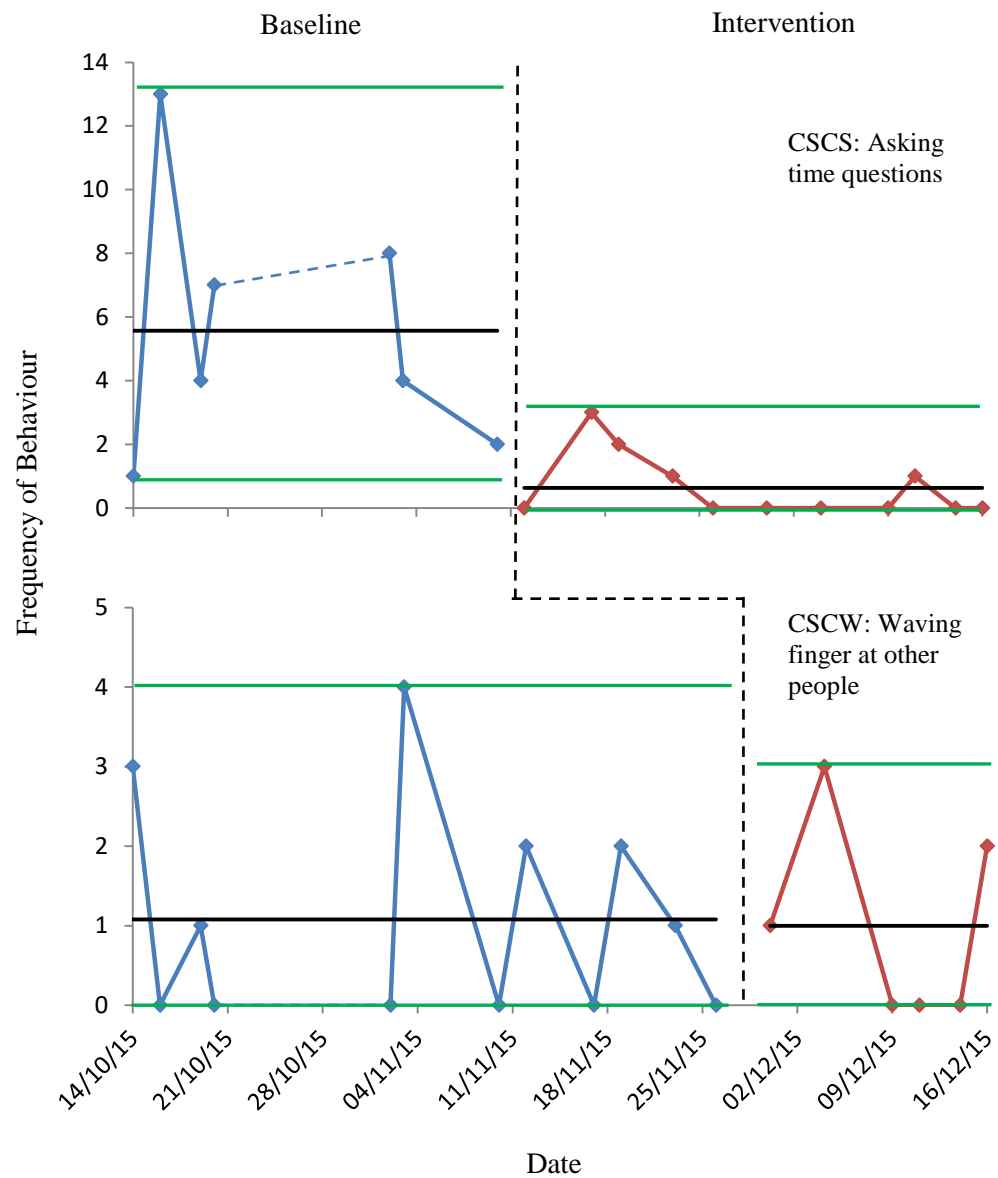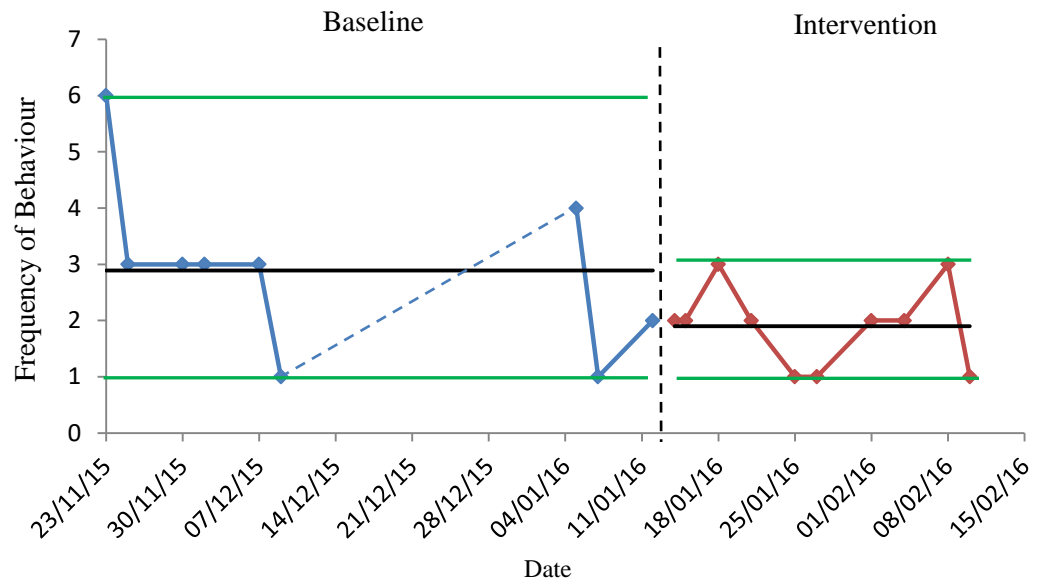*Graph 4.9: A scattergraph showing the frequency of Robert's target behaviours across baseline and intervention phases with mean lines and variability lines*

Characteristic	Description of outcome for TB1
Level	Mean of baseline phase = 2.89 Mean of intervention phase = 1.9 There was a mean level decrease of 0.99 between baseline and intervention.
Trend	Graph 4.8 shows a downward trend in the baseline phase and then a less steep downward trend in the intervention phase.
Variability	Graph 4.9 shows a moderate level of variability of scores in the baseline phase (range = 5; S.D = 1.54). There is less variability in the intervention phase (range = 2; S.D = 0.74).
Immediacy of Effect	There was no immediate change of level between the final three data points in the baseline phase and the first three in the intervention phase.
Overlap	All (100%) of the data in the intervention phase overlapped with those in the baseline phase.

Table 4.10: Summary of the outcome of visual analysis for Robert's frequency of behaviours graphs

4.5.3.3   Results of Tau-U Analysis

Behaviour	Tau-U Effect Size	Confidence Interval (90%)	P-Value
1 (CSCS)	-0.27	-0.71 < > 0.18	0.33

Table 4.11: Results of the Tau-U effect size analysis with confidence intervals and p-values (Robert)

4.5.3.4   Summary of Outcomes of Repeated Measures

TB1: CSCS

The visual analysis suggests a minimal intervention effect. There is a small decrease in the mean (-0.99) between phases and less variability in the intervention phase which may indicate a small effect. However, there is a downward trend during the baseline phase, suggesting that some improvement

110

could be predicted regardless of the intervention, although the unstable baseline makes analysis more difficult. The 100% data overlap additionally moderates any assertion of an intervention effect.

In concordance with the visual analysis, the Tau-U analysis produced an effect size of -0.27, which can be classed as a 'small' effect and this was not a statistically significant result ($p = 0.33$).

4.5.3.5 Results of Pre- and Post-Intervention Measure

TB1: CSCS

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	7	9
Post-intervention	2	2
Difference	-5	-7

Table 4.12: Staff member ratings given on the pre- and post-intervention questionnaires for TB1 (Robert)

Table 4.12 shows that there was a moderate decrease in the staff member's rating of Robert's TB1 in relation to how challenging/disruptive they find it to be and a larger decrease in how much of an impact they perceive it to have on his learning and/or social relationships. This data is not supportive of the results of the visual analysis and effect size analysis, which identified only a small possible effect.

The staff member rated the CSCW intervention at '7' for effectiveness in addressing TB1 and commented that Robert:

'Responded to this intervention in a really positive way. He was able to gain an understanding of why/how he should respond when given a task. This has helped him to avoid disruptive behaviour and encouraged him to settle."

4.5.4   Case Study Four: Owen

4.5.4.1   Pupil Profile

Age at start of data collection: 5 years and 9 months

TB1: Calling out 'No'

TB2: Interfering with peers' use of toys/objects during free-choice time

Order of conditions: Owen started with the CSCW intervention (targeting TB1) followed by CSCS (targeting TB2).

Number of CSCs completed in CSCW condition: 4

4.5.4.2　　　Outcomes of Repeated Measures
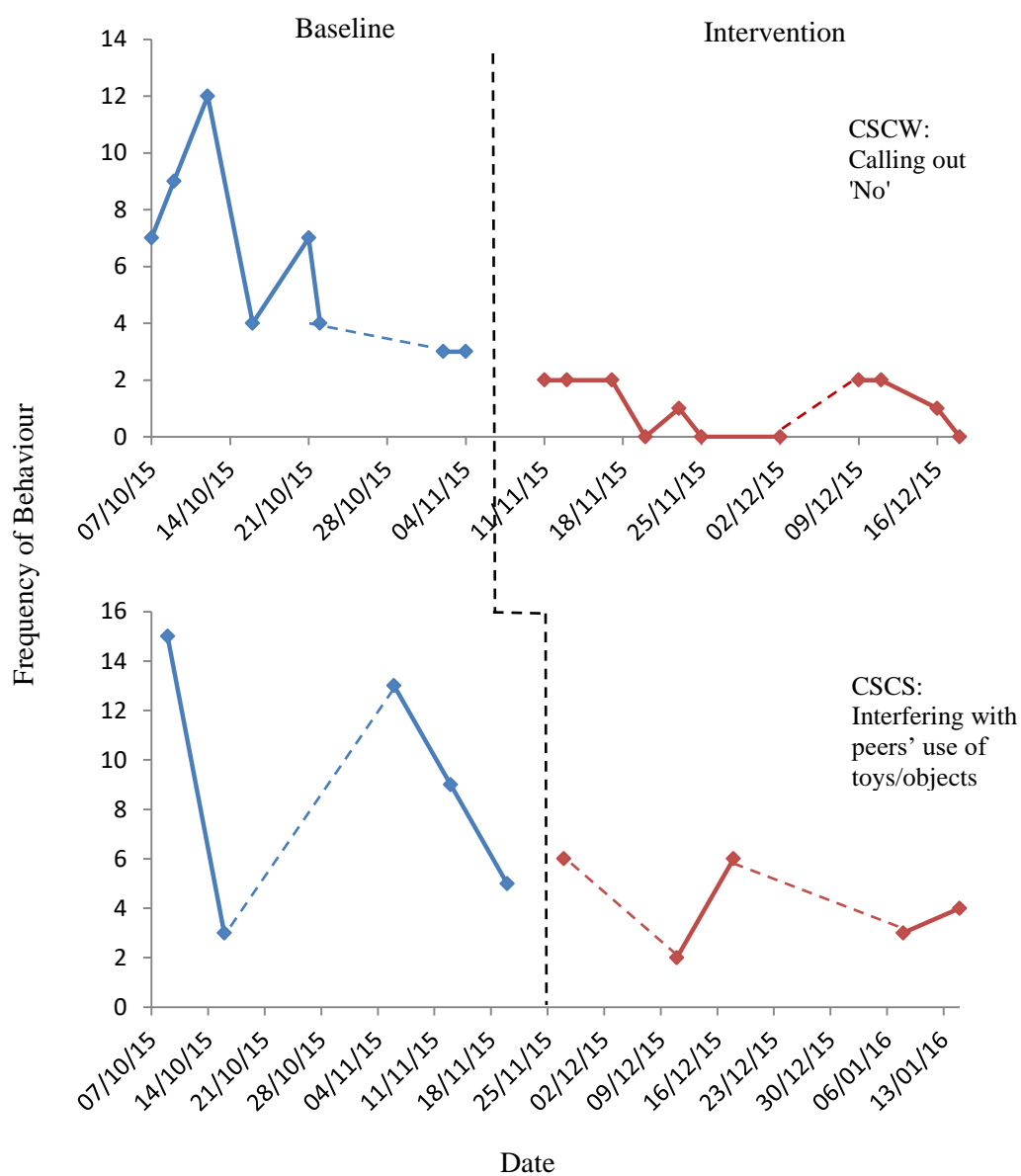
Graph 4.10: A scattergraph showing the frequency of Owen's target behaviours across baseline and intervention phases
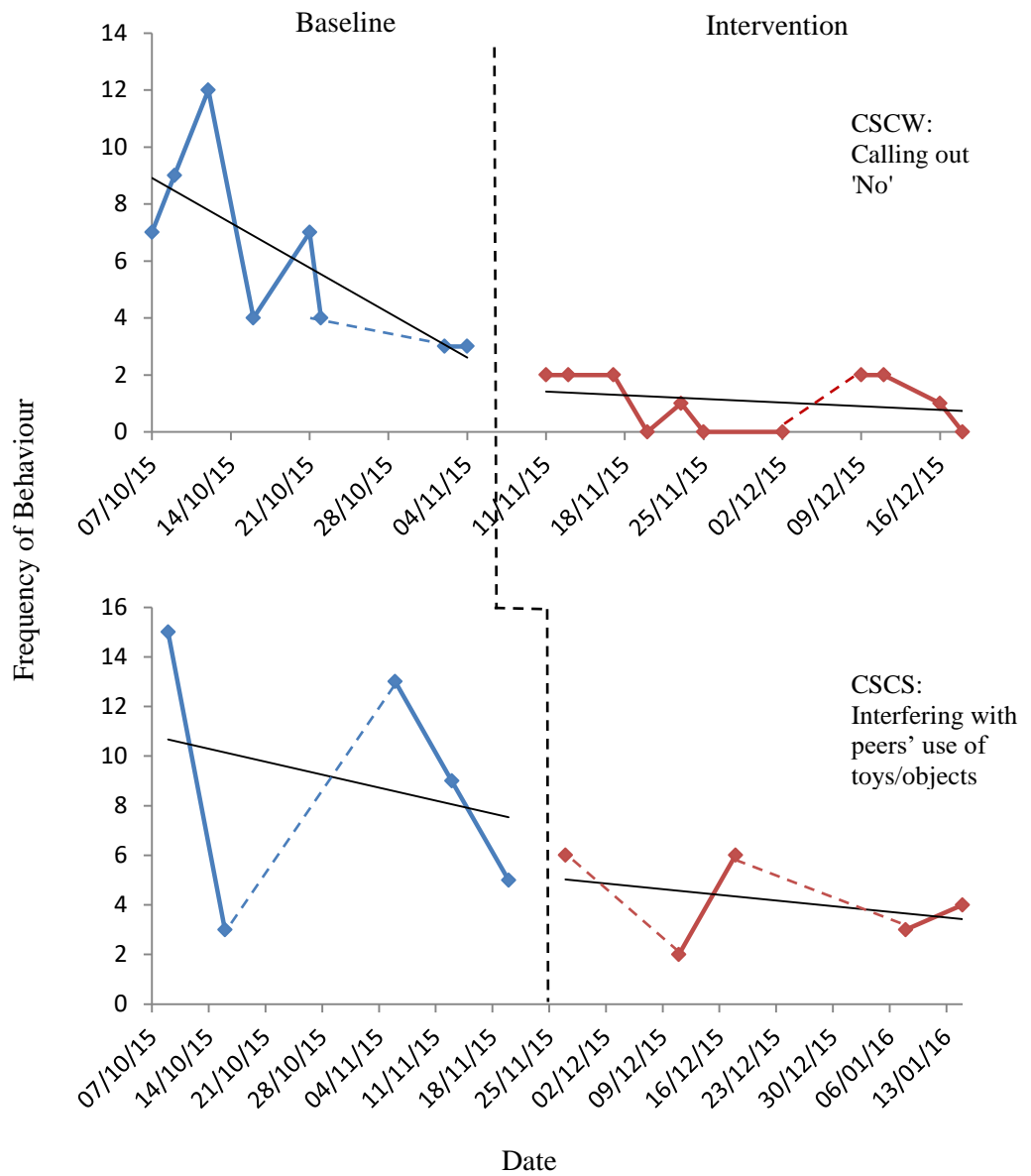
Graph 4.11: A scattergraph showing the frequency of Owen's target behaviours across baseline and intervention phases with trendlines
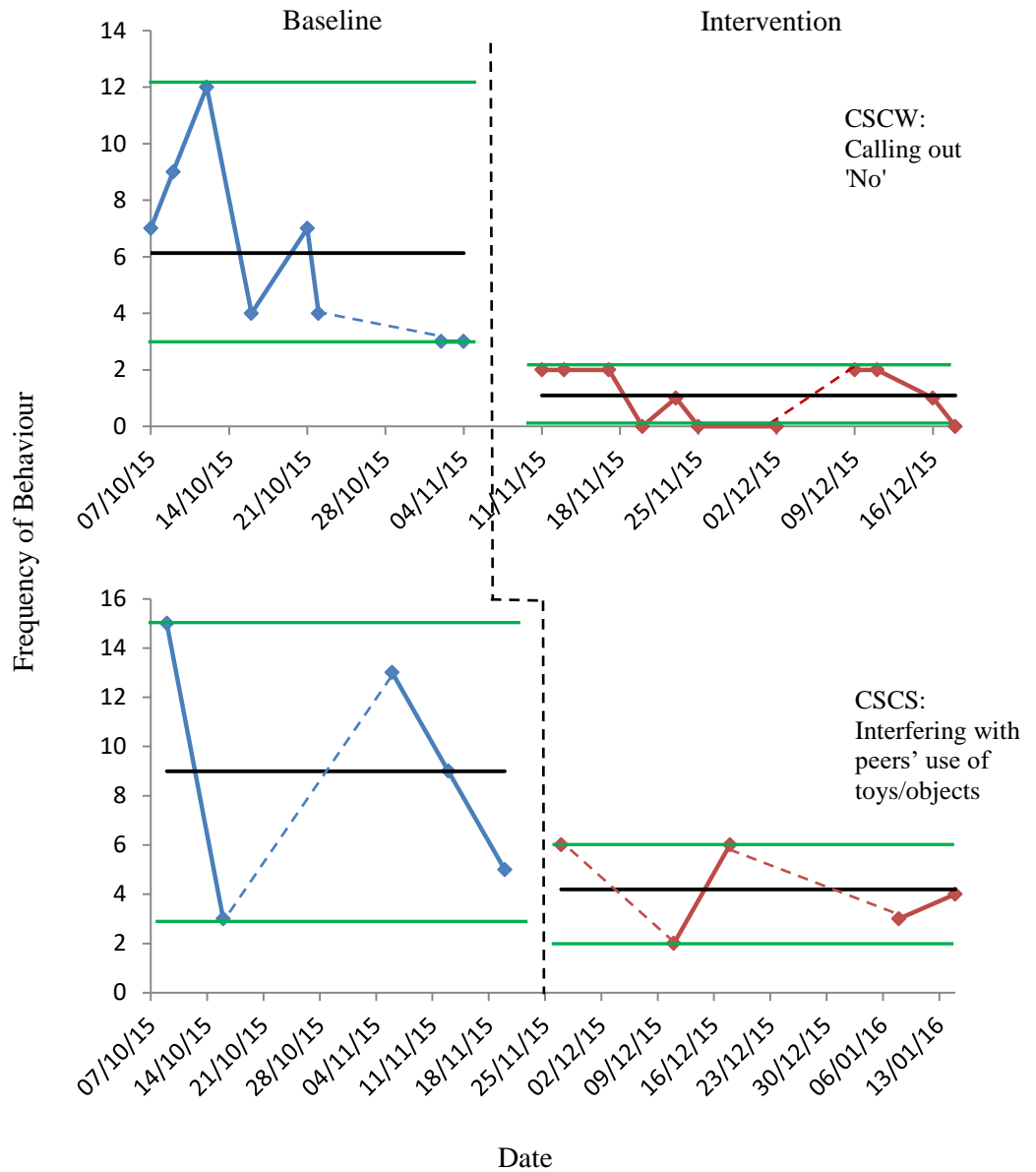
Graph 4.12: A scattergraph showing the frequency of Owen's target behaviours across baseline and intervention phases with mean lines and variability lines

Characteristic	Description of outcome for TB1	Description of outcome for TB2
Level	Mean of baseline phase = 6.13 Mean of intervention phase = 1.09 There was a mean level decrease of 5.04 between baseline and intervention.	Mean of baseline phase = 9 Mean of intervention phase = 4.2 There was a mean level decrease of 4.8 between baseline and intervention.
Trend	Graph 4.11 shows a clear downward trend in the baseline phase. There is little trend in the intervention phase.	Graph 4.11 shows a slight downward trend in the baseline phase and a similar level of downward slope in the intervention phase.
Variability	Graph 4.12 shows a fairly high level of variability of scores in the baseline phase (range = 9; S.D = 3.23). There is much less variability in the intervention phase (range = 2; S.D = 0.94).	Graph 4.12 shows a high level of variability in the baseline phase (range = 12; S.D = 5.10). There is much less variability in the intervention phase (range = 4; S.D = 1.79)
Immediacy of Effect	There was an immediate (but not marked) change in level between the final three data points in the baseline phase and the first three in the intervention phase.	There was an overall decrease in level between the last three data points in the baseline phase and the first three in the intervention phase.
Overlap	None (0%) of the data points in the intervention phase overlapped with those in the baseline phase.	80% of the data points in the intervention phase overlapped with those in the baseline phase.

Table 4.13: Summary of the outcome of visual analysis for Owen's frequency of behaviours graphs

4.5.4.3 Results of Tau-U Analysis

Behaviour	Tau-U Effect Size	Confidence Interval (90%)	P-Value
1 (CSCW)	-0.81	-1.26 < > -0.35	< 0.01
2 (CSCS)	-0.40	-1.03 < > 0.23	0.30

Table 4.14: Results of the Tau-U effect size analysis with confidence intervals and p-values (Owen)

4.5.4.4 Summary of Outcomes of Repeated Measures

TB1: CSCW

The visual analysis suggests a positive intervention effect. There is a notable decrease (-5.04) in the mean between phases, no data overlap, and much less variability in the intervention phase. The assertion of a possible intervention effect is confounded somewhat by the unstable baseline which also displays a downward trend, suggesting that some improvement could be predicted regardless of the intervention.

However, the Tau-U analysis (which was able to control for baseline trend) produced an effect size of -0.81, which can be described as a 'large' effect and this was a statistically significant finding ($p = <0.01$).

TB2: CSCS

The visual analysis is suggestive of a possible intervention effect. There is a decrease in the mean (-4.8) between phases and much lower levels of variability in the intervention phase. However, the unstable baseline makes analysis more difficult, and there exists a downward trend during baseline suggesting that some improvement could be predicted regardless of the intervention. There is also a high proportion of data overlap (80%) between phases which moderates any assertion of an intervention effect. In addition, missing data makes analysing patterns difficult.

The Tau-U analysis produced an effect size of -0.40 which can be classed as a 'small' effect. This finding was not statistically significant ($p = 0.30$).

In considering the multiple-baseline graph as a whole, this does not reliably demonstrate that a change occurred only when the intervention was directed at the target behaviour, as there was a concurrent decrease in the frequency of TB2 at the time of intervention implementation for TB1, although the generally unstable baselines make identifying clear changes very difficult.

4.5.4.5 Results of Pre- and Post-Intervention Measures

TB1: CSCW

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	8	7
Post-intervention	4	3
Difference	-4	-4

Table 4.15: Staff member ratings given on the pre- and post-intervention questionnaires for TB1 (Owen)

Table 4.15 shows that there was a moderate decrease in the staff member's rating of Owen's TB1 in relation to how challenging/disruptive they find it to be and how much of an impact it has on his learning and/or social relationships. This data is generally supportive of the results of the visual analysis and effect size analysis, although perhaps is not reflective of the 'large' effect size produced by the Tau-U analysis.

The staff member rated the CSCW intervention at '7' for effectiveness in addressing TB1 and commented that Owen:

"Enjoyed doing the comic strips, especially the drawings" and that she felt he was *"understanding more not to say 'No' but also why this is sometimes inappropriate"*.

TB2: CSCS

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	5	8
Post-intervention	5	8
Difference	0	0

Table 4.16: Staff member ratings given on the pre- and post-intervention questionnaires for TB2 (Owen)

Table 4.16 shows that there was no change in the staff member's rating of Owen's TB2 in relation to how challenging/disruptive they find it to be and how much of an impact it has on his learning and/or social relationships. This data does not support the results of the visual analysis or effect size analysis, which were indicative of a possible modest effect.

Despite the lack of change in the above ratings, the staff member rated the CSCS at '6' for effectiveness and commented:

"It was clear that the intervention had less of an effect on this behaviour. I feel that may be due to the fact that it was only addressed once" and that Owen had *"found it difficult to relate to it (the CSC) when the situation was happening."*

4.5.5   Case Study Five: Gareth

4.5.5.1   Pupil Profile

Age at start of data collection: 8 years and 6 months

TB1: Calling out

TB2: Touching/taking items belonging to others

Order of conditions: Gareth started with the CSCS intervention (targeting TB1) followed by CSCW (targeting TB2).

Number of CSCs completed in CSCW condition: 3

4.5.5.2   Outcomes of Repeated Measures
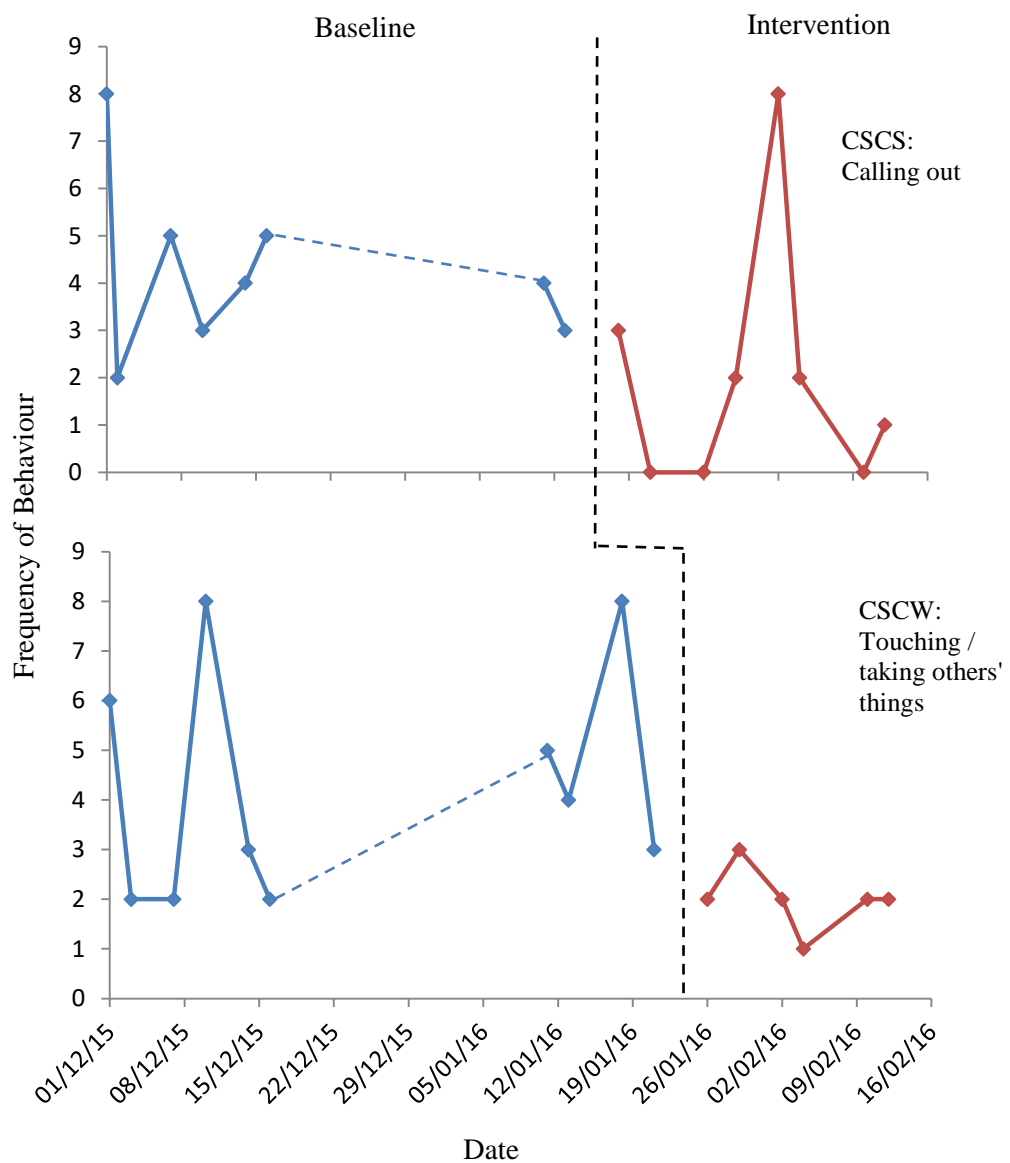
Graph 4.13: A scattergraph showing the frequency of Gareth's target behaviours across baseline and intervention phases
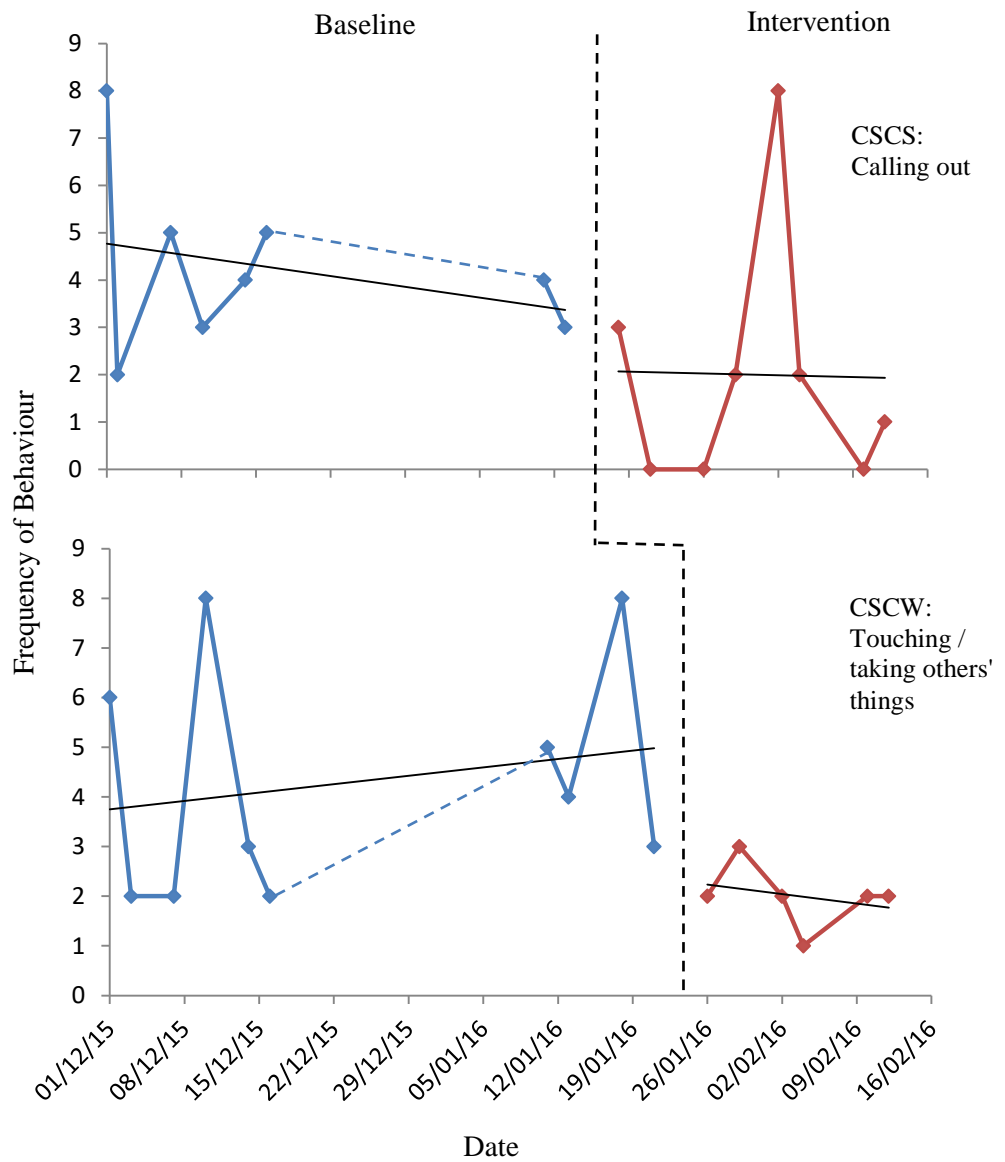
Graph 4.14: A scattergraph showing the frequency of Gareth's target behaviours across baseline and intervention phases with trendlines
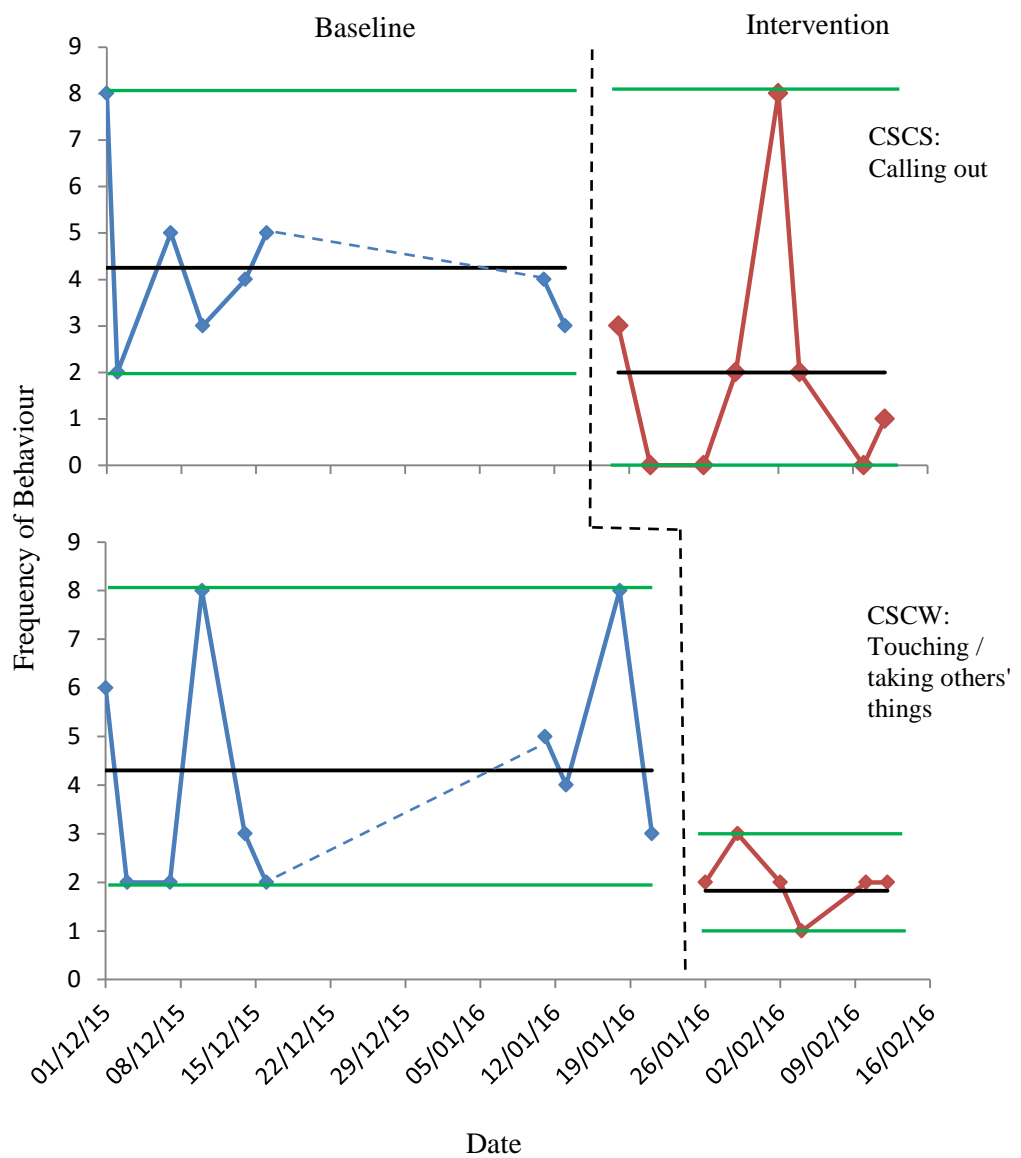
Graph 4.15: A scattergraph showing the frequency of Gareth's target behaviours across baseline and intervention phases with mean lines and variability lines

Characteristic	Description of outcome for TB1	Description of outcome for TB2
Level	Mean of baseline phase = 4.25 Mean of intervention phase = 2 There was a mean level decrease of 2.25 between baseline and intervention.	Mean of baseline phase = 4.3 Mean of intervention phase = 2 There was a mean level decrease of 2.3 between baseline and intervention.
Trend	Graph 4.14 shows slight downward trend in the baseline phase. There is a very slight downward in the intervention phase.	Graph 4.14 shows a slight upward trend in the baseline phase. There is a downward trend in the intervention phase.
Variability	Graph 4.15 shows some variability of scores in the baseline phase (range = 6; S.D = 1.83). There is a higher level of variability in the intervention phase (range = 8; S.D = 2.75).	Graph 4.15 shows some variability in the baseline phase (range = 6; S.D = 2.36). There is less variability in the intervention phase (range = 2; S.D = 0.75).
Immediacy of Effect	There was an overall decrease in level between the last three data points in the baseline phase and the first three in the intervention phase.	There was an overall decrease in level between the last three data points in the baseline phase and the first three in the intervention phase.
Overlap	50% of the data points in the intervention phase overlapped with those in the baseline phase.	83% of the data points in the intervention phase overlapped with those in the baseline phase.

Table 4.17: Summary of the outcome of visual analysis for Gareth's frequency of behaviours graphs

4.5.5.3 Results of Tau-U Analysis

Behaviour	Tau-U Effect Size	Confidence Interval (90%)	P-Value
1 (CSCS)	-0.59	-1.08 < > -0.10	< 0.05
2 (CSCW)	-0.77	-1.27 < > -0.26	0.01

Table 4.18: Results of the Tau-U effect size analysis with confidence intervals and p-values (Gareth)

4.5.5.4 Summary of Outcomes of Repeated Measures

TB1: CSCS

The visual analysis is suggestive of a possible intervention effect. This is suggested by the modest decrease in the mean (-2.25) between phases, a fairly immediate effect after the start of the intervention phase and only a moderate proportion of data overlap (50%). However, an unstable baseline makes analysis more difficult and the downward trend during baseline (which is not maintained during the intervention phase) suggests that some improvement could be predicted regardless of the intervention. In addition, a higher level of variability within the intervention phase leads to questions about the reliability of this data, although this is mainly due to one particular outlier.

The Tau-U analysis produced an effect size of -0.59, which can be classed as a 'moderate' effect size. This was a statistically significant result ($p < 0.05$).

TB2: CSCW

The visual analysis is suggestive of a possible intervention effect. There is a small decrease in the mean (-2.3) between phases and less variability in the intervention phase. There is also a slight upward trend during baseline which suggests that an increase in the behaviour could be predicted to continue in the absence of an intervention. Once the intervention was introduced, there was an immediate effect and the intervention phase data displays a desirable, downward trend. Factors that moderate the assertion of an intervention effect are the unstable baseline which makes analysis more difficult and the high level of data overlap (83%).

The Tau-U analysis produced an effect size of -0.77, which can be classed as a 'moderate' effect. This was a statistically significant result (p = 0.01).

In considering the multiple-baseline graph as a whole, there is a small amount of evidence that change occurred only when the intervention was directed at the target behaviour, as there was an increase in TB2 following the implementation of the intervention for TB1, although this did then decrease just before the intervention was implemented for TB2. However, the generally unstable baselines make identifying clear changes very difficult.

4.5.5.5   Results of Pre- and Post-Intervention Measures

TB1: CSCS

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	6	6
Post-intervention	2	2
Difference	-4	-4

Table 4.19: Staff member ratings given on the pre- and post-intervention questionnaires for TB1 (Gareth)

Table 4.19 shows that there was a moderate decrease in the staff member's rating of Gareth's TB1 in relation to how challenging/disruptive they find it to be and how much of an impact it has on his learning and/or social relationships. This data is supportive of the results of the visual analysis and effect size analysis.

The staff member rated the CSCS intervention at '5' for effectiveness in addressing TB1 and commented that Gareth:

"Was reluctant to do the CSCs at first but as time went on he was more willing" and that he had *"enjoyed drawing the CSC but was not very keen on discussing feelings of others as he found it difficult to put himself in other people's positions."*

TB2: CSCW

	How challenging/disruptive is behaviour perceived to be?	How much of an impact does the behaviour have on learning/social relationships?
Pre-intervention	8	7
Post-intervention	5	4
Difference	-3	-3

Table 4.20: Staff member ratings given on the pre- and post-intervention questionnaires for TB2 (Gareth)

Table 4.20 shows that there was a modest decrease in the staff member's rating of Gareth's TB2 in relation to how challenging/disruptive they find it to be and how much of an impact it has on his learning and/or social relationships. The decrease was less than that for TB1. This data is not fully supportive of the results of the visual analysis and effect size analysis, the latter of which found an effect size of -0.77 (larger than for TB1).

The staff member rated the CSCW intervention at '6' for effectiveness in addressing TB2. They provided no additional comments.

4.6   Research Question 2: Summary of Outcomes

Research Question 2 asked:

'Does the level of improvement in target behaviours vary in relation to the frequency with which the CSC intervention is implemented?'

Table 4.21 displays the effect size measure for both intervention conditions for each participant. Based on the effect size measure, it appears that the CSCW intervention was more effective for three out of the four participants who were exposed to both conditions, although the difference was only marginal in Jack's case.

Participant	Intervention Condition	Tau-U Effect Size Outcome	Most Effective Condition
Jack	CSCS	-0.22	CSCW
	CSCW	-0.29	
Daniel	CSCS	-0.88	CSCS
	CSCW	-0.01	
Robert	CSCS	-0.27	N/A
Owen	CSCS	-0.40	CSCW
	CSCW	-0.81	
Gareth	CSCS	-0.59	CSCW
	CSCW	-0.77	

Table 4.21: The relative effectiveness of the CSCS and CSCW intervention for each participant (measured by Tau-U effect size)

4.7 Intervention Fidelity

As detailed in section 3.9.10, the researcher conducted intervention fidelity checks through three methods:

1. Staff completion of CSC intervention fidelity checklists

2. Examination of completed CSCs by the researcher

3. Examination of intervention diaries

Intervention Fidelity Checklist Data:

Examination of the checklists indicated that intervention fidelity levels were 100% for the majority of the CSCs completed and this was confirmed through examination of the completed CSCs. However, there were some notable exceptions to the 100% fidelity rates which need to be borne in mind when interpreting the outcomes of the intervention, as outlined below. These shall be further discussed in Chapter 5.

- The CSC that Jack completed for TB2 (CSCS condition) did not include a recorded action plan. The intervention diary indicates that possible action plans were discussed verbally during some review sessions but these were never recorded onto his CSC.

- One of Jack's CSCs for TB1 (CSCW condition) lacked speech and thought bubbles and lacked a clear sequence.

- The CSC that Gareth created for TB1 (CSCS condition) did not include a recorded action plan. The staff member commented that they verbally suggested some possibilities at the time and during reviews, but these were not recorded onto the CSC.

Intervention Diary Data:

As already indicated, the planned intervention schedule for Daniel was not adhered to which resulted in him only completing two CSCs during the CSCW intervention phase. In Gareth's case, there was a slight delay in completing the

first CSC (CSCS) due to the staff member experiencing a lack of co-operation from him to begin with.

The intervention diaries also highlighted that in all cases regular, brief reviews of completed CSCs were conducted to support the participants in keeping their CSCs in mind during the course of the research.

4.8   Inter-Observer Reliability

As discussed previously in section 3.10.3.2, some observations were conducted jointly by a staff member and the researcher in order to assess IOA levels and inform reliability of the repeated measures. The researcher aimed to achieve a minimum of 15% joint observation sessions and this was achieved in most cases. Exceptions to this were the observations of Owen's TB2, for which only 10% were conducted jointly due to lack of availability of staff, and Robert, for whom no joint observations were carried out as a result of time pressures meaning that the researcher was unable to spend a whole morning at the school.

Following the guidelines provided by Friman (2009), IOA levels were calculated in the following ways:

- for event sample data, the smaller figure was divided by the larger figure, then multiplied by 100 to get a percentage.

- for interval sample data, the number of intervals with agreement between observers was divided by the total number of intervals, then multiplied by 100 to get a percentage.

Table 4.22 displays the results of these calculations.

Participant	Target Behaviour	% of joint observations conducted (to 1 d.p)	Overall % of agreement (to 1 d.p)
Jack	TB1	14.3%	95%
	TB2	14.3%	100%
Daniel	TB1	16.7%	100%
	TB2	16.7%	91.6%
Robert	TB1	0%	N/A
Owen	TB1	15.8%	96.3%
	TB2	10%	90%
Gareth	TB1	18.8%	87.5%
	TB2	18.8%	94.4%

Table 4.22: The percentage of joint observations conducted for each participant and the associated levels of inter-observer agreement

Guidelines recommend that the minimum acceptable level of percentage agreement is 80%, and 90% is preferred (Friman, 2009; Kratochwill et al, 2013). Based on these criteria, all meet the minimum level of 80% agreement and all but one meet the higher preferred figure of 90%. This enhances confidence in the reliability of the observational data.

4.9 Inter-Rater Reliability of Visual Analysis

In an attempt to address the issue of subjectivity in visual analysis, the researcher and another trainee educational psychologist (who was familiar with the visual analysis of SCED data) separately analysed each graph using the

131

visual analysis guidance outlined in section 4.2. Although effect size analysis was used to complement the visual analysis and aid in identifying possible effects that may have been visually obscured due to variable data sets, the researcher still deemed it worthwhile to collect inter-rater agreement data with regards to the visual analysis. The researcher was keen that the utility of visual analysis not be dismissed in the face of standardised numerical outcomes and recognised the limitations of context-free effect size analysis.

Both raters had access to the three graphs for each participant displaying the raw data only, raw data with trendlines, and raw data with mean lines and variability lines. They could also view the visual analysis summary tables. See Appendix 17 for a copy of the script used in the inter-rater assessment.

With regards to this information, and in line with the recommendations of Brossart et al (2006), the raters then rated the following statement on a scale of 1 (not at all certain) to 5 (very certain):

'How certain or convinced are you that there was a practical, significant improvement in the participant's behaviour between baseline and intervention phase?'

The level of agreement between raters was statistically analysed using the Cohen's Kappa statistic (Cohen, 1960) (see Appendix 18 for the Kappa output). According to guidelines from Altman (1999), a Kappa of <0.20 indicates 'poor' agreement, 0.21-0.40 indicates 'fair' agreement, 0.41-0.60 indicates 'moderate agreement', 0.61-0.80 indicates 'good' agreement, and 0.81-1.00 indicates 'very good' agreement.

The level of agreement between raters was 0.52, indicating a 'moderate' level of agreement. Achieving only a moderate level of inter-rater agreement provides further justification for including additional effect size analysis of the data.

Chapter 5 - Discussion

5.1 Introduction to Chapter

This chapter aims to outline and further explore the results presented in Chapter 4. The chapter begins by summarising the findings in relation to each research question, and interpreting these where possible in relation to the key theory and research introduced in Chapter 2 (Literature Review) and the design and methodological procedures presented in Chapter 3 (Methodology).

Limitations of the current research are discussed, as are implications for practice and suggestions for future research. The chapter concludes with the reflections of the researcher on the research experience.

5.2 Summary of Findings: Research Questions 1a and 1b

Research questions 1a and 1b asked:

1a) Do CSCs have a positive impact on the target behaviours of primary-aged pupils on the autistic spectrum?

1b) Are the findings of the repeated measures reflected in perceived change in these target behaviours from the perspective of school staff?

5.2.1 Jack

Visual analysis provided limited evidence of an intervention effect for both target behaviours. This was reflected in the Tau-U effect size analysis which indicated only 'small' intervention effects. The multiple-baseline graph did not reliably demonstrate that a change occurred only when the intervention was directed at a target behaviour.

In reasonable concordance with the repeated measures analysis, staff perceptions of Jack's behaviours changed either very little or not at all between pre- and post-assessment. However, they rated the effectiveness of the CSCS and CSCW interventions at a moderate '6' for effectiveness, which seems

somewhat in contrast with their other ratings. Further discussion of issues relating to the pre- and post-intervention measure can be found in section 5.8.1.5.

5.2.2 Daniel

Visual analysis was indicative of an intervention effect for TB1 which was supported by a 'large' and statistically significant effect size of -0.88. Both visual analysis and effect size analysis concluded no intervention effect for TB2, although baseline data was already at, or near to, the floor prior to implementing the intervention and it also became apparent that the CSCW intervention had not been delivered to the intended schedule. The multiple-baseline graph did not reliably demonstrate that a change occurred only when the intervention was directed at a target behaviour.

The large effect size for TB1 was reflected in the findings of the pre-and post-intervention measure and the rating of '9' given to the effectiveness of the intervention. As would be expected given the different outcome for TB2, there was a more modest change in staff ratings for this behaviour and a rating of '6' was given for effectiveness. However, these appeared higher than would be expected given the negligible change identified by the repeated measures analysis.

5.2.3 Robert

Visual analysis concluded a minimal intervention effect which was confirmed by a 'small' effect size of -0.27. The change in ratings given by staff on the pre- and post-intervention measure, and an effectiveness rating of '7', were greater than would have been predicted by the outcomes of the repeated measures.

5.2.4 Owen

Visual analysis was indicative of an intervention effect for TB1 which was supported by a large and statistically significant effect size of -0.81. Visual

analysis suggested possible evidence of an intervention effect for TB2 which was supported by an effect size of -0.40. However, missing data was a problem in the data set for TB2 meaning this data may be unreliable and not fully reflect true patterns of behaviour. The multiple-baseline graph did not reliably demonstrate that a change occurred only when the intervention was directed at a target behaviour.

The large effect size for TB1 was reflected to a reasonable degree in the change in ratings between pre-and post-measures of staff perceptions of the behaviour and the rating of '7' given to the effectiveness of the intervention. There was no change in the ratings of staff perceptions of TB2, despite an effect size of 0.40. However, in contrast to the other ratings, a rating of '6' was given for effectiveness.

5.2.5   Gareth

Visual analysis was indicative of a possible intervention effect for TB1 which was confirmed by a 'moderate' and statistically significant effect size of -0.59. Visual analysis of the data for TB2 again suggested a possible intervention effect and another 'moderate' and statistically significant effect size of -0.77 was obtained. The multiple-baseline graph indicated some evidence that meaningful change occurred only when the intervention was introduced for each behaviour, thus strengthening the assertion of a possible intervention effect for this participant.

The pre- and post- intervention questionnaire showed some concordance with the repeated measures analysis in that there was a decrease in staff ratings for both TB1 and TB2, and respective effectiveness ratings of '5' and '6' were given.

5.3   Overall Summary of Findings for Research Questions 1a and 1b:

The summary of findings for each participant highlights the ideographic approach that this research adopted in its investigation of the impact of the

CSC intervention on five individuals. This section now considers overall themes in the findings across cases.

5.3.1   Research Question 1a:

In addition to visual analysis, Tau-U analyses were conducted to add a level of validation and scrutiny to the results. By this measure, effect sizes greater than that which could be called trivial (i.e. >0.20) (Cohen, 1988) were obtained across all but one of the behaviours addressed through CSCs in this research. However, there was significant variability in the effect sizes obtained between participants and no clear pattern of responding emerged between participants.

One participant (Gareth) appeared to respond well to both CSC interventions, reflected in the 'moderate' and statistically significant effect sizes for both behaviours. Two participants (Daniel and Owen) appeared to respond well to one of the CSC interventions, as reflected by the 'large' and statistically significant effect sizes for one of their behaviours, but only small, non-significant effect sizes for the other behaviour. Two participants (Jack and Robert) showed lower levels of response as reflected in the 'small' and non-significant effect sizes for their behaviours.

For ease of comparison, the benchmarks proposed by Cohen (1988) and Ferguson (2009) were used when interpreting effect sizes. The dangers of using terms like 'small', 'moderate' or 'large' out of context have been outlined in section 4.4.1. What is of most importance is the *practical* significance of an effect, dependent on its relative costs and benefits. Although not extensively examined in this research, efforts were made to gauge staff perceptions of the impact of the CSC intervention on the target behaviours. In some cases, this showed that some behaviours were felt by staff to be having less of an impact on learning and/or social relationships following the intervention, even where repeated measures analysis suggested otherwise (e.g. Robert's TB1 and Daniel's TB2). The researcher would argue that this could be important information to take into account when considering the practical significance of results.

It is important to note that the confidence intervals obtained for all the participants' effect size results were relatively large, even for those results which reached the level of statistical significance. This is likely a result of the relatively small data sets and it therefore must be acknowledged that the effect sizes obtained may have low reliability, even those which appear large and visually convincing (Parker et al, 2007). Whilst statistically significant results can bolster the assertion of an intervention effect, it is important to remember that there is a danger of drawing firm conclusions where the sample is too small to justify such confidence as there is a heightened risk of a type I error.

As such, it is acknowledged that the present results can only lead to tentative, speculative conclusions as certain design issues and data characteristics (to be explored further in future sections) mean causal inference cannot be clearly established and therefore the experimental hypothesis cannot be accepted with confidence. The results appear to indicate that CSCs show some promise as an effective intervention, even when used for relatively short periods, although not necessarily in all cases. Further discussion around this can be found in later sections.

5.3.2   Research Question 1b:

As with the visual and effect size analyses, the data obtained from the pre- and post-intervention measure varied between participants. Practical changes in ratings related to how disruptive or challenging staff perceived the behaviour to be, and how much of an impact it was felt the behaviour had on learning and/or social relationships, were indicated across all the behaviours targeted in three cases (Daniel, Robert and Gareth), across just one target behaviour in one case (Owen), and across neither behaviour in one case (Jack). Staff ratings of the effectiveness of the interventions ranged from five to nine. In some cases, the pre- and post-intervention data did not triangulate particularly closely with the repeated measure analysis. Possible reasons for this are discussed in section 5.5.

5.4   Interpretation of Findings: Research Questions 1a

The findings of this research indicated positive results for some participants, although there were different patterns of responding and inconsistent outcomes both between and within participants. In considering this finding in relation to the existing literature on CSCs, previous research has reported positive outcomes for most, if not all, of the participants involved. However, one factor in particular to consider is that some previous studies have only involved one participant (Glaeser et al, 2003; Rogers & Myles, 2001; Vivian et al, 2012). It seems plausible to predict that different levels of response may have been identified if more participants had been involved, as was the case in this research.

A number of factors make direct comparison with much of the previous research difficult. These include differences in design, as some previous research has not been experimental (Glaeser et al, 2003; Pierson & Glaeser, 2007; Rogers & Myles, 2001; Vivian et al, 2012), and differences in data analysis procedures. In addition, diagnoses need to be considered when comparing the results of research as participants have been variously described as having ASC, high-functioning autism, or Asperger's syndrome. Differences in relation to the specifics of diagnoses and in the conceptualisation of ASC could impact on the outcomes of research. One participant in this research had an additional diagnosis of ADHD (Robert) which adds another layer of complexity when comparing results across participants in different studies.

There exist only a very limited number of experimental studies into the effectiveness of CSCs to which to compare the results of this research. The findings of this research show similarities to those of Ahmed-Husain & Dunsmuir (2014), who also used a multiple-baseline design to investigate the effectiveness of CSCs in addressing target behaviours of secondary-aged pupils with diagnoses of ASC. These researchers also used Tau-U effect size analysis in addition to visual analysis and examination of these indicated that, as in the present research, effect sizes ranged from small to large both within and between participants. Although not experimental in design, the results of the study by Pierson & Glaeser (2007) found the target behaviours of three

primary-aged pupils with 'high-functioning autism' to have reduced by 50 – 75% over a six week intervention period, again indicating different levels of effect between participants.

In this research, for some of the participants there was an apparent intervention effect within a relatively short space of time. In the cases where less improvement was seen, it is possible that a longer intervention period may have resulted in a greater impact, as it has been suggested in previous research that the effects of CSCs may increase after a period of five to six weeks (Pierson & Glaeser, 2007). However, there has been a variety of intervention durations used in previous studies, ranging from four weeks to several months, so at present it remains unclear as to how rapidly one might expect to see changes. It may well be that, as with other factors, this is likely to vary depending on the nature of the behaviour and other characteristics of the individual and their context.

5.4.1.1   Theoretical Interpretations

As outlined in Chapter 2, the theory of mind hypothesis is a key psychological theory informing understanding of the social communication and interaction difficulties and differences in autism (Baron-Cohen et al, 1985). CSCs involve creating visual systems which are designed to support understanding of the more abstract elements of interaction, such as thoughts and feelings. In this way, CSCs aim to support the 'mentalising' skills of individuals with autism and support the social interaction, communication and adaptation difficulties that are the most defining feature of the condition (Jones, 2002). Indeed, some research has indicated gains in theory of mind ability following a CSC intervention (Lewandowski et al, 2014; Vivian et al, 2012). Although it was beyond the scope of this research to explicitly explore the process and mechanisms by which the CSCs may have induced change (as the focus was on outcome evaluation), a number of additional comments provided by the staff involved indicate that support for theory of mind development was perhaps evident:

"Using the CSC was good for putting the participant in the shoes of the children involved and to talk about how they might be feeling." (Jack)

"He was able to gain an understanding of how and <u>why</u> he should respond when given a task."(Robert)

"I think he is understanding more not to say 'No' but also <u>why</u> this is sometimes inappropriate." (Owen)

However, this process was not always easy as evidenced in this quote about Gareth:

"(He) was not very keen on discussing the feelings of others as he found it difficult to put himself in other people's positions"

This final quote demonstrates why the development of interventions designed to support this area of difficulty is so important.

5.4.1.2 Use of Visual Strategies

The use of visual strategies and cues can help improve social interaction and understanding in individuals with autism due to the tendency of this population to be able to process visual information more easily than verbal information (e.g. Kerr & Durkin, 2004; Theimann and Goldstein, 2001; Quill, 1997). The highly visual nature of CSCs is supportive of this as it reduces the reliance on verbal language. This may help to explain why some of the participants appeared to respond well to CSCs after not previously responding well to purely verbal prompts and requests to refrain from certain behaviours. Some of the additional comments from staff indicated that participants appreciated the highly visual nature of the intervention*:*

"(He) was open to illustrating problems" (Daniel)

"(He) enjoyed doing the comic strips, especially the drawings." (Owen)

5.4.1.3   Nature of Target Behaviours

A range of behaviours have been targeted in this, and previous, research. The researcher proposes that differences in outcomes could be mediated by the nature of the particular behaviour targeted, such as how engrained or complex they are. For example, Daniels TB1 (asking time-related questions) was reported to have started to become an issue during the summer holidays, with this continuing once back at school in September. As such, it was not a particularly long-standing behaviour which may have made it more amenable to change. In other cases, behaviours were reported to have been more longstanding. It is possible that a longer intervention period may have been needed to achieve an identifiable impact on such behaviours.

5.4.1.4   CSC procedures

As explained in section 4.7 (and to be further explored in section 5.8.2), intervention fidelity was lacking for some of the CSCs created by Jack and Gareth. It is important to consider this when interpreting the results for these participants, especially as only small effect sizes were obtained for both of Jack's behaviours. Interestingly, a moderate effect size was found for Gareth's TB1, despite not visually recording an action plan for this behaviour. However, possible alternative actions were reportedly discussed verbally (as indicated by the intervention diary). This finding links with those of Ahmed-Husain & Dunsmuir (2014) who found CSCs can be effective with auditory, rather than visual, action plans in some cases, and that this may relate to an individual's relative strengths in either verbal or visual processing.

The extent to which colour was used varied between the CSCs created by the five participants. All but Gareth used colour to some extent in some, if not all, of their CSCs. It is possible that this feature of CSCs plays a significant role in determining its effectiveness, due to the additional understanding that using colour as a visual representation of the feelings of the self and others may provide. Such variety in actual CSC content was anticipated to some extent, as each CSC is unique to that individual, and Gray (1994b) suggested that colours were not an immediately necessary component but could be gradually

introduced as appropriate over the course of several CSCs. In the present findings, there was not a clear link indicated between the use of colour in CSCs and intervention effectiveness, as positive outcomes were indicted for Gareth despite him being the only participant whose CSCs were absent of this feature.

5.4.1.5 CSC reviews

Participants were provided with opportunities to review and discuss previously completed CSCs. Intervention diary comments indicate that Daniel, Robert and Owen had added to their action plans through the process of review. Whilst this level of flexibility was deemed important in terms of participants maintaining an active sense of ownership of their CSCs (Vivian et al, 2012), it is possible that it may have been these reviews, rather than the original process of creating the CSC, that made more of an impact.

5.4.1.6 Individual Differences

Other researchers have begun to query the possible mediating effect of cognitive ability on the effectiveness of CSCs (Ahmed-Husain & Dunsmuir, 2014; Vivian et al, 20102). Although not explicitly assessed in this research, some basic information about attainments in school was gathered and this indicated that there were a range of abilities amongst the participants. However, the results of this research do not necessarily indicate that those with higher ability responded better to the intervention, as the two participants for who 'large' effects were found were known to be working considerably below age expectations.

Although not explored in detail, additional comments included in the intervention diaries indicated that participants' engagement levels varied. Diary comments indicated that Daniel, Robert and Owen generally engaged very well and enjoyed completing their comic strips. Jack's level of motivation and interest appeared to be quite low during some of the intervention period, perhaps reflected in his apparent lower levels of response. The staff member working with Gareth also expressed difficulties in achieving motivation and engagement at times. However, there were indications of greater impact on

Gareth's behaviours than Jack's despite these difficulties. Again, this perhaps reflects a complex picture of the way that CSCs may interact with participant characteristics and behaviours to have an effect.

5.4.1.7 Conclusions

The results of the repeated measures analysis in this research provide some support to the findings of previous research into the effectiveness of CSCs. The researcher would agree with those who have suggested that a number of factors, such as context and participant characteristics, may interact with the CSC topic and content, thus resulting in uneven results for CSC interventions (Lewandowski et al, 2014). This is in line with conclusions drawn from meta-analyses of other story-based interventions (such as Social Stories) that 'the most striking feature of the data…is the degree of inconsistency' (Reynhout & Carter, 2006, p.466). One needs to keep in mind the heterogeneous nature of the autistic population, and observations that no one intervention has been found to be effective for every individual with autism (Diehl, 2003).

Although the present findings are indicative of the potential effectiveness of CSCs in at least some cases, there are a number of methodological issues which could lead to alternative explanations for the findings. These will be discussed later in the chapter.

5.5 Interpretation of Findings: Research Question 1b

It was not the aim to establish causal relationships with the data collected via the pre- and post-intervention questionnaire, but rather to use it for triangulation purposes. In the cases in which there were apparent differences between this data and the repeated measures analysis, this may have been a reflection of the various threats to validity and reliability of data collected at single points in time that were recognised in section 3.10.3.3. Due to these threats, it is possibly not surprising that the findings were sometimes in contrast to the repeated measures data that attempted to control for such threats. A possible explanation for the apparently inflated 'effectiveness' ratings in some cases (namely those for Jack's two TBs, Daniel's TB2 and

Robert's TB1) may be respondent bias, as staff were not blind to the aims of the research and may have given higher scores to satisfy the researcher.

It is also important to remember that this measure was subjective in nature and therefore a certain level of change in behaviour in one participant may have led to a different magnitude in change ratings than it would have for another participant, due to factors such as staff expectations of behaviour and tolerance levels.

5.6 Summary of Findings - Research Question 2

Research question 2 asked:

'Does the level of improvement in target behaviours vary in relation to the frequency with which the CSC intervention is implemented?'

The Tau-U effect size analysis was able to identify, in a standardised manner, possible differential levels of intervention impact as a result of different frequencies of intervention implementation. The results of this indicated that the CSCW intervention was more effective than the CSCS intervention for three out of the four participants who were exposed to both conditions, although this was only a marginal difference in one case (Jack).

The CSC effectiveness ratings that staff provided on the post-intervention questionnaire can also be considered as a triangulation measure in relation to this research question, and these were consistent with the effect size analysis. The CSCW intervention was rated as more effective than the CSCS intervention for Owen and Gareth. The two conditions were rated as equally effective for Jack, in concordance with the similar effect sizes obtained for both behaviours in this case. Daniel's CSCS intervention received a higher effectiveness rating than the CSCW intervention, as reflected in the effect sizes for this participant. However, issues relating to the pre- and post-intervention data have already been outlined so it would not be appropriate to place too much weight on these additional findings.

In summary, there is some indication that, in some cases, completing more regular CSCs may result in greater improvement in behaviours than if only a

single CSC is developed. However, findings related to this research question were not consistent and no firm conclusions can be drawn at this point from the small sample involved in this research (i.e. the experimental hypothesis cannot be accepted).

5.7 Interpretation of Findings: Research Question 2

The findings in relation to this question require some consideration as this is the first piece of research to explicitly investigate this. Previous research has indicated positive results across a wide range of intervention frequencies (although design issues need to be considered in many of these). Ahmed-Husain & Dunsmuir (2014) found positive results with just one CSC, whereas other studies have included much higher numbers over the course the intervention period (e.g. Glaeser et al, 2003; Lewandowski et al, 2014; Vivian et al, 2012).

Whilst the present findings are clearly limited due to the small sample size, there is some early indication that increased frequency may secure better outcomes. The exception to this finding was Daniel, where the CSCW condition implied almost no effect. However, as previously stated, the frequency of the behaviour targeted in this condition was already at or near the floor for much of the baseline data set, leaving little room for improvement to be seen anyway, and only two CSCs were actually completed.

Although additional comments on the perceived relative effectiveness of each intervention condition were not explicitly sought, the staff member who implemented the intervention with Owen commented in relation to the CSCS for TB2 that:

"It was clear that the intervention had less of an effect on this behaviour...I feel that may be due to the fact that it was only addressed once" and that Owen had *"found it difficult to relate to it (the CSC) when the situation was happening"*.

These comments may reflect that, in some cases at least, opportunities to consider their behaviour across a number of different situations could lead to

greater impacts. This can be linked to possible implications of weak central coherence (Frith, 1989) for the education of children with ASC, in that they should not necessarily be expected to generalise previous learning in new situations without training or specific prompting (Frederickson & Cline, 2009).

The researcher considers that this idea could also be linked to the tendency of individuals with autism to adhere to rules once they have been internalised (Scattone et al, 2002). With CSCs, the clear social information provided and the formation of a visually recorded action plan could help to reduce the frequency of target behaviours through providing individuals with a concrete representation or 'rule' of what they should do in a specific situation. In cases where the target behaviour is relatively straightforward and unambiguous, it may be that a relevant 'rule' can be adequately captured within one CSC. For more complex, entrenched or multi-faceted behaviours, a number of CSCs focusing on different instances of the behaviour may be required to provide that individual with greater opportunity to understand the relevant situations and for them to develop effective means of responding differently in the future.

Indeed, it has already been suggested that the number of CSCs needed to establish positive outcomes is likely to vary significantly across individuals, targets and contexts (Gray, 1998; Vivian et al, 2012). The present findings appear to support this suggestion as one participant displayed a very positive response after only one CSC whereas other participants displayed more significant improvements when a series of CSCs were completed.

5.8 Evaluation of the Research Methodology

Through visual and effect size analysis, the present findings indicated practically significant intervention effects in some cases. However, an evaluation of research design is imperative before conclusions can confidently be drawn (Kratochwill et al, 2013).

5.8.1 Research Design

This research used a multiple-baseline SCED to investigate the impact of CSCs on target behaviours. This type of design can enhance internal validity as if there are changes in the target to which the intervention is applied, but not to other targets at that time, there is a stronger case for arguing causal relationships (Robson, 2011). For one participant (Robert), an AB design was used and this significantly restricts the ability to infer any causal relation in this case.

A number of limitations remain with regards to the design and mean that the results of this research need to be considered within certain parameters.

5.8.1.1 Number of Demonstrations of an Effect

In their criteria for SCED standards, Kratochwill et al (2013) argue that a study should demonstrate at least three temporally distinct indications of an intervention effect before a researcher can confidently assert that there is sufficient evidence of an intervention's effectiveness. In this research, there were only two opportunities to consider possible intervention effects as only two behaviours were targeted and no 'return to baseline' phase was included. The researcher would contest that the inclusion of multiple participants can compensate to some extent for this as an apparent replication of an intervention effect across a number of participants helps to support the argument of an intervention's effectiveness.

Visual analysis of the multiple-baseline graphs indicted that, in most cases, there was no clear evidence that changes in behaviour occurred only in

response to the intervention, due to the irregular baseline patterns present in most of the data sets. This reduces the internal validity of the results as it means that observed effects cannot be confidently attributed to the intervention itself.

5.8.1.2 Stability of Baselines

As in all SCEDs, the stability of the baseline is highly important. This could be said to be especially so in multiple-baseline design, where one wishes to be able to identify a clear effect when, and only when, an intervention is introduced. It is difficult to infer causal relationships if baseline data does not provide sufficient demonstration of a clearly defined pattern of responding that can be used to extrapolate predicted future performance, assuming no changes to the independent variable (Barlow et al, 2009; Kratochwill et al, 2010). Within this research, baselines were limited for both ethical and practical reasons and stable baselines were not established for any of the participants, which is often the case in real world applied research (Robson, 2011). However, this was acknowledged throughout the analysis and is recognised as a limitation of the present research.

Baseline trend is another factor to consider. Noticeable downward trend was observed in the baseline data for a number of target behaviours. These trends were problematic as they were in the direction of the predicted effect of the intervention. However, the Tau-U analysis provided a valuable tool to address this, and correcting for baseline trend led to more conservative effect sizes than would have been obtained if these trends were ignored.

5.8.1.3 Length of Phases

The length of phases was largely determined by the time constraints within which the researcher was working. Within all phases for each participant, there were at least five data points, which is the recommended minimum (Horner et al, 2005), and in most cases there were more than this, which allowed for a longer period of time in which patterns of responding could be identified.

As would be expected in a multiple-baseline design, the length of phases varied. The intervention phases for some of the second target behaviours (namely Jack, Daniel and Gareth) were slightly shorter than anticipated due to delays in staff commencing the intervention for these behaviours. This could have had an impact on results as there was less time for the intervention to have an effect. However, in the cases where possible intervention effects were evidenced, these seemed to occur quite quickly after the introduction of the intervention, as opposed to showing delayed effects, which goes some way to alleviating concerns that larger effects may in some cases simply be the result of longer intervention periods. The measure taken to counterbalance the order of conditions between participants also helped to alleviate this potential difficulty.

5.8.1.4 Observational Measures

Direct observational measures are argued to increase the validity of SCED data (Kratochwill et al, 1992), but the observation schedules used in this research needed to account for a range of possible sources of bias or error (see Table 3.9). One key measure taken to address issues of reliability of the observational data was to conduct some observations jointly with a staff member in order to calculate IOA, and the outcomes of this showed high levels of agreement of between 87.5% and 100%. It is important to note that only a small percentage (10%) of observations were conducted jointly for Owen's TB2 and none were conducted for Robert for the reasons outlined in section 4.8. It therefore must be acknowledged that there is less certainty about the reliability of this data which should be taken into account when considering their results. Additionally, the researcher acknowledges that a higher proportion of joint observations across all cases would have allowed for greater confidence in the reliability of the observational measures.

It is possible that the presence of a second observer affected the performance of the first observer (e.g. they may have been more conscientious with their recording). Therefore, one cannot rule out the possibility that, in the context of a busy classroom, frequency recordings were always accurate.

Another important aspect informing the validity and reliability of the observation measures is the need for them to be operationally defined (Horner et al 2005; Robson, 2011). The researcher endeavoured to work alongside staff to develop operational definitions of behaviour (see Table *3.6*). The inclusion of a trial period of data collection allowed for ambiguities to be highlighted and definitions to be amended if necessary, and no further queries regarding the definitions of the target behaviours were made following this. This is reflected in the high IOA scores where these are available.

5.8.1.5 Pre- and Post-Intervention Measure

The pre- and post-intervention measure collected data at only two points in time, meaning there was no opportunity to examine patterns over time for this variable. As such, the potential impact of other factors, not related to the CSC intervention, need to be considered where changes in this measure were identified. As highlighted in section 3.10.3.3, the possibility of respondent bias and a high level of subjectivity also mean that the findings of this measure should be treated with caution.

Despite these difficulties, the researcher would argue that staff perceptions are highly relevant to the evaluation of interventions in real life school contexts. Therefore, future research may benefit from consideration of how to gain these in as valid and reliable a way as possible.

5.8.1.6 Missing data

On the whole, data was collected according to the agreed schedule of twice per week. Where holiday periods fell during baseline phases, the researcher ensured that at least a couple of further baseline data points were gathered after the holiday period in order to give time for things to settle again before introducing the intervention. A few data points were missed in some cases due to pupil or staff absence but this was not deemed to have a significant impact on results as, in general, sufficiently regular observations were conducted for the purposes of identifying patterns in the data. An exception to this was the

data collection for Owen's TB2, which experienced a fair amount of disruption and this needs to be taken into account when interpreting the results for this behaviour, especially as data collection was limited to once per week at best for this behaviour.

5.8.2 Intervention Fidelity

As outlined in section 4.7, high levels of intervention fidelity were indicated in most cases. However, there were some notable exceptions which need to be borne in mind when considering outcomes.

One of Jack's CSCs for TB1 (CSCW condition) lacked speech and thought bubbles. These are deemed a critical aspect of the intervention as they provide visual support to help the pupil understand the more abstract elements of social interaction, in particular the thoughts and feelings of others. The CSC that Jack completed for TB2 (CSCS condition) did not include a recorded action plan. Although the intervention diary indicated that possible action plans were discussed verbally during some review sessions, the lack of a visual record of these could have impacted on outcomes. It is important to bear these factors in mind, especially in view of the results for Jack which indicated only a small possible intervention effect for both behaviours.

The CSC that Gareth created for TB1 (CSCS condition) did not include a recorded action plan. The staff member commented that they verbally suggested some possibilities at the time and during reviews, but these were not recorded onto the CSC. Again, this should be considered when interpreting the results for this behaviour (although generally positive outcomes were indicated regardless of this).

5.8.3 Analysis of Data

Preliminary data analysis was completed through visual analysis, following the criteria of Kratochwill et al (2013), to help identify whether practical changes were evident. The limitations of visual analysis in terms of subjectivity were outlined in section 4.2.1, and therefore efforts were made to assess the

reliability of this through calculating inter-rater agreement using Cohen's Kappa. This indicated a 'moderate' level of agreement.

It is now widely advocated that single case researchers should supplement visual analysis with an effect size measure in order to provide more standardised and reliable results, and to assist in the identification of results that may not be large and obvious, yet still may be of practical significance (Brossart et al, 2014; Kratochwill et al, 2013; Parker et al, 2007; Vannest & Ninci, 2015). As outlined in section 4.4.1, Tau-U effect size analysis was used as this was deemed a highly appropriate option given certain characteristics of the data (e.g. baseline trend and short data sets). This provided a standardised measure of outcomes, which was particularly pertinent to Research Question 2, and allowed for a more fine-tuned analysis in terms of considering the relative impact of the intervention across participants.

Whilst the researcher attempted a comprehensive analysis of the data, the limitations of both the visual analysis, in respect of the difficulties posed by undesirable data characteristics such as high variability, and the effect size analysis, in terms of its low reliability (see section 5.3.1) and inability to take into account contextual factors, are acknowledged.

5.9 Internal Validity

As with other fixed designs, SCEDs are at risk from a variety of threats to internal validity, as outlined in section 3.10.1. There are a number of threats to internal validity that are deemed particularly pertinent to this research, as outlined below.

5.9.1 History

The researcher was informed that Jack had begun receiving weekly 'positive play' sessions during the baseline data collection phase. These sessions were said to be targeting skills such as listening to instructions, sharing and making appropriate eye contact. It is possible that the targets of the 'positive play'

intervention showed generalisation to the behaviours of interest in this research and could have resulted in any behaviour change identified (even if small).

In the case of Gareth, there was a change to his typical teaching environment after the Christmas holiday in that he was spending less time within small group sessions outside of his classroom and more time within the whole class. This change in context could have impacted on the frequency of his behaviours.

It is also important to note that the researcher did not attempt to control for the use of additional prompts or positive verbal feedback from staff during the research, therefore it is feasible that it was factors such as these that led to behaviour change, in the cases where that was evidenced, rather than the actual process of creating the CSCs. However, verbal feedback such as this is used anyway as part of everyday practice in schools, and staff will instinctively do this in order to promote the well-being, development and motivation of their pupils. In terms of maintaining ecological validity, it was deemed important to maintain as naturalistic an environment as possible but the potential threat to validity that this presents should be taken into account.

An additional history-related threat is that, as mentioned previously, due to the research design and time-scales involved, some school holidays fell within the data collection period. This was only a one-week half term in some cases (Jack, Daniel and Owen's TB1), but a longer two-week Christmas holiday during the data collection for Robert, Owen's TB2 and Gareth. The researcher acknowledges that this is an extraneous variable which may have impacted on results and these holiday periods have been indicated on the graphs to ensure transparency in the data. Additionally, the staff member working with Owen commented to the researcher after the October half-term break that his mother had noticed how often he was saying 'No' at home and had been talking to him about this. This obviously needs to be considered when interpreting the outcomes of this behaviour, which appeared to be very positive.

5.9.2 Maturation

The inclusion of multiple cases can help to reduce this threat if intervention effects are demonstrated across different participants. There was some evidence of this but not consistently. In addition, as there was not clear evidence in the multiple-baseline graphs of an intervention effect only when the intervention was introduced for a particular behaviour, this remains a threat to the validity of the results. However, as the research period was relatively short, it would not be expected that maturation would have as large an impact as it may in studies of longer duration.

5.9.3 Hawthorne Effect

Blinding procedures were not used in this research as the researcher wanted to ensure informed consent from all the participants, in line with ethical standards. This means that the possibility that behaviour change occurred due to participants being aware of the research focus, rather than the manipulation of the independent variable, cannot be ruled out.

5.10 Implications for EP Practice

The research conducted to date indicates that there is a slowly growing evidence base providing some support for the use of CSCs in addressing behaviours of concern in children with ASC. The outcomes of this study suggest that the implementation of a CSC intervention by staff in a primary school setting may have resulted in practical changes in behaviour for some pupils. However, findings to date have not been unequivocal and design limitations of the present and previous research should not be overlooked. The researcher would, however, propose that there is sufficient promising preliminary evidence to warrant the intervention being researched further.

This research has indicated that CSCs are a relatively straightforward intervention for school staff to deliver, requiring no monetary expense and not being very demanding on time, especially if positive outcomes are secured following the completion of a single CSC. The use of a highly visual structure

such as this will be familiar to many school staff who are often aware that the use of visual strategies are advocated as an element of good practice in the teaching of pupils with autism (Jones, 2006; Mesibov & Howley, 2003). These factors are likely to increase the appeal of the intervention to school staff. EPs would be well-placed to provide the necessary training to school staff to enable them to understand the theory behind CSCs and how to implement them.

However, it will be important for EPs to communicate to school staff and parents that the evidence base for CSCs is still in its infancy. Providing an honest summary of existing research will enable staff to make more informed decisions about whether or not they wish to try the intervention. In terms of guidance on the frequency with which to implement a CSC intervention, the research is even more limited, with the present research being the only attempt so far to explicitly investigate this. However, given the results of this research, along with the positive findings of previous research involving a range of CSC frequencies, it appears reasonable to suggest to staff that consideration of factors such as pupil characteristics, the nature of the behaviour(s) of concern and the range of contexts in which it occurs may help to inform planning around how often to implement the intervention. As with all interventions, close monitoring of a pupil's response should be the key factor which guides intervention planning and delivery.

Based on the findings of the present and previous research into CSCs, the researcher has developed a 'decision tree' (see page 158) which could potentially be utilised by school staff and other professionals (e.g. EPs, AOT) to help guide their considerations about whether or not a CSC intervention may be appropriate to use with a specific pupil. This has taken into account the suggestions of a number of authors (e.g. Gray, 1994b; Vivian et al, 2012; Ahmed-Husain & Dunsmuir, 2014; Lewandowski et al, 2014) who have hypothesised that the impact of CSCs may perhaps be moderated by factors such as cognitive ability, the nature of the behaviour being targeted, and the context(s) in which it occurs. The mixed results of the present research also appear to support such possibilities.

The decision tree incorporates initial key questions relating to certain factors that may be of relevance to the successful implementation of the intervention, and then further questions encourage more in-depth deliberation about the characteristics of the pupil, the behaviour of concern, and the staff/setting supporting them. It is intended that all three areas be considered and then the answers to the questions, plus any additional discussion that emerges, may help to illuminate how appropriate a CSC intervention may be for that pupil.

It is important to note that, given the current limited evidence base for CSCs, the factors included in the decision tree reflect preliminary ideas and are not based on firm empirical data. As such, it is not intended to be a definitive guide to whether or not CSCs should be used in a given situation, and it is unlikely that there will be simple 'Yes' or 'No' answers to a number of the questions posed for consideration. The guide is rather intended for use as a supportive and reflective tool which could be drawn upon when considering how best to support a pupil's development, alongside consideration of alternative or complementary intervention strategies.

CSC 'Decision Tree'

Pupil Factors

Does the pupil have sufficient cognitive/verbal ability to be able to access the intervention and take part in discussions?

If yes, also consider...

Would the pupil be able to actively engage with an adult, on a 1:1 basis?

Do you feel the pupil would be motivated by the style of the intervention? *(e.g. consider their interest and strengths in the use of visual techniques such as drawings/symbols).*

Nature of Behaviour

Does the behaviour appear suitable to be targeted by a CSC intervention? (e.g. *consider if it lends itself to the discussion of the thoughts and feelings of self and others).*

If yes, also consider...

Is the behaviour deemed relatively straightforward to address - *consider how many contexts it occurs/how complex it is (i.e. could it be sufficiently captured within a short, simple CSC?)*

Is the behaviour of concern one that has emerged relatively recently? *More entrenched behaviours may require a longer or more intense intervention period, and/or alternative intervention*

Practical/Setting Factors

Would staff availability allow dedicated time to hold 1:1 sessions with the pupil to construct and review CSCs?

If yes, also consider...

Consider the supporting adult's knowledge of the pupil and their skills in engaging with them, including a commitment to encouraging active participation and 2-way learning.

Do you feel you would require more information or training about the use of CSCs before implementing this intervention?

If 'Yes' has been answered to most questions and no significant obstacles have been identified, it may indicate that CSCs could be a useful intervention to try.

157

There appears a further implication with regards to how the effectiveness of this (or indeed any) intervention or strategy is assessed in daily EP practice. The use of subjective ratings by staff in this research highlighted that outcomes of these can sometimes be at odds with more objective behavioural data. The collection of rating scale data to establish the effectiveness of strategies and interventions, such as through a 'target monitoring evaluation (TME)' framework (Dunsmuir et al, 2009), is a common practice used in the educational psychology service within which the researcher was on placement. Whilst staff perceptions are indeed important, if one hopes to acquire more objective information to inform 'practice-based evidence' (Fox, 2011), then the potential limitations of perception scales should be considered.

In order to increase EPs confidence in recommending the approach, as well as other professionals such as Autism Outreach staff, future research of the type outlined in section 5.11 will be needed. With the growing drive for evidence-based practice, practitioners should endeavour to ensure that the approaches and interventions they recommend are drawn from an adequate research base.

EPs could have an important role to play in contributing to the research base for CSCs in their role as 'scientist-practitioners'. The limitations of RCTs (the traditional 'gold standard') in terms of gaining an understanding of the relative impact of interventions on individuals have been acknowledged (Frederickson, 2002; Harrington, 2001; Neef, 2009). The findings of the present research support this idea and provide further support for the use of SCEDs in intervention evaluation research, and such designs can realistically be implemented within EP practice (Ali & Frederickson, 2006).

5.11 Future Research

The nature of the methodology employed in this research necessarily limits the extent to which findings can be generalised. This limitation can be addressed through replication (Horner et al, 2005), so the present research would therefore benefit from replication across a larger number of cases and across different settings. However, the researcher would recommend certain adaptations to be made to the methodology in order for future research to demonstrate increased internal validity. These include:

- the use of a design which allows for three temporally distinct demonstrations of an effect to be considered, as per the guidance of Kratochwill et al (2013), for example through including three separate behaviours or using a multiple-baseline across participants design in which at least three participants are involved.

- aiming to achieve more stable baselines before introducing the intervention, although the difficulty in doing this is appreciated.

- aiming to ensure a high level of intervention fidelity across all cases as this was a limitation in some cases in this research.

Due to time constraints, maintenance of intervention effects was not assessed in the present research. Some previous experimental research has included assessment of maintenance (Ahmed-Husain & Dunsmuir, 2014; Lewandowski et al, 2014) and the inclusion of this in future research would be valuable in determining more clearly whether positive outcomes can be expected to extend into the longer-term.

In adopting a post-positivist perspective, the present research based its conclusions on quantitative data. Future research may be able to extend understanding of the context in which interventions are delivered and the impact that contextual factors may have on outcomes through the collection of supplementary qualitative data (Burden, 2015). For example, this could take the form of interviews or focus groups and could include both staff and pupils.

Dedicated assessment of the views of participants about the CSC intervention has not been a feature of research to date and could be a worthwhile avenue to explore in the future.

A further avenue to explore could be the key components of CSCs, for example, through investigating the relative impacts of the discussion around the behaviours and mental states of the self and others, the use of pictorial representations of these (e.g. thought bubbles), and the impact of the use of colours as representations of emotional states. However, the researcher would suggest that the accumulation of an evidence base for the general effectiveness of CSCs should be an initial key objective, alongside the possible deeper exploration of the component processes.

Finally, when considering the practical significance of results of intervention research, one needs to look in context of the change that is desired or expected. In terms of CSCs, this is difficult at present as there is so little research which has reported effect sizes, therefore there is not yet a clear picture of what 'typical' effects could be expected. Future research may benefit from embracing the advances that have been made in regard to calculating effect sizes in SCEDs in order to see if effect sizes reported thus far can be replicated, and aiding any future meta-analysis of CSC research.

5.12 Researcher's Reflections

The researcher appreciates the valuable and rewarding experience that she has had in planning and implementing a research study within real-life school contexts. However, this was not without its challenges and has highlighted the myriad of factors that cannot be easily (if at all) controlled within such research settings.

Some key practical difficulties experienced by the researcher included the initial recruitment of participants and establishing the means of obtaining regular measures, the latter being heavily influenced by the availability of support staff and their already busy timetables within schools. This highlighted the importance of engaging and communicating effectively with stakeholders.

The researcher also experienced the difficulties that can arise in conducting and drawing conclusions from research involving children with ASC due to their diverse range of characteristics. This required the researcher to give careful consideration to the design and measures used to try and make them as relevant and applicable as possible within the practical and logistical parameters of this research.

Intervention infidelity and lack of adhering to intervention schedules was an issue that occurred in some cases, despite regular communication. This led the researcher to reflect on the initial information sharing and training sessions with staff which, with hindsight, may have been enhanced by communicating more detailed information about the process and rationale of the research. However, as stated previously, whilst it was intended that the intervention be delivered following standard guidelines as far as possible, the researcher appreciated that the participants (and also the staff delivering the intervention) had their own unique characteristics, strengths and needs which would inevitably impact on the course of the intervention which is, by its nature, a dynamic one in which two individuals are engaged in discussion about a unique set of circumstances.

The researcher hoped to make a useful contribution to the currently limited evidence base for CSCs. The researcher has certainly gained increased knowledge in this area which will usefully inform future practice. Despite its recognised limitations and inconclusive findings, the researcher reflects that this is an area worthy of future study to help inform the ever-evolving landscape of educational support for children and young people with ASC.

A final reflection relates to the researcher's enhanced appreciation of EPs as 'scientist-practitioners', being well placed to use their research skills within their practice. The use of SCEDs, such as those used in this research, appear to lend themselves well to the work that EPs carry out at the individual child level.

Chapter 6 -   Conclusions

6.1   Main Findings

This research evaluated the effectiveness of CSCs for improving the target behaviours of primary-aged pupils with autism in mainstream school settings. Analysis of the SCEDs showed a high level of variability in outcomes across participants and clear patterns of responding following the introduction of the intervention were not established. In summary, one participant appeared to respond well to both CSC interventions, two participants appeared to respond well to one of the CSC interventions and two participants showed lower levels of response.

The pre- and post-intervention triangulation measure showed practical improvements in staff perceptions of target behaviours for six out of the nine total behaviours of interest in this research. In some cases, changes in ratings between pre- and post-assessment appeared to triangulate well or adequately well with the repeated measures outcomes but this was not consistent across all cases. Post-intervention ratings of the effectiveness of the intervention for addressing the target behaviours varied and, again, showed differing levels of agreement with the repeated measures outcomes. However, the threats to the validity and reliability of single-point data outlined in the discussion restrict any causal conclusions that can be drawn from these findings, which is why they were included as a triangulation measure only.

In terms of answering Research Question 2, analysis suggested that the CSCW intervention was more effective than the CSCS intervention for three out of four participants, although there was only a marginal difference in one case. It is considered that no firm conclusions can be drawn at this point and it may be the case that the number of CSCs needed to establish positive outcomes will vary depending on the characteristics of individuals, behaviours and contexts.

When considering outcomes, it is imperative to acknowledge several limitations of the research. One key issue was the unstable baseline data sets which made visual identification of intervention effects more difficult. This

issue was alleviated somewhat through the addition of an appropriate effect size measure, although large-confidence intervals around the obtained effect sizes meant that these findings could not be deemed to be highly reliable.

A multiple-baseline design was utilised in order to enhance the internal validity of the findings through being able to assess if there were causal links between the intervention and behaviour change across targets. These were not clearly established and therefore left open the possibility that threats to internal validity could offer alternative explanations for the results. Particularly relevant to this research were the threats of history and the Hawthorne effect. Missing data and intervention fidelity issues in some cases further complicated the interpretation of results.

Despite its limitations, this research illustrates a potential positive impact of the CSC intervention for some pupils with ASC and further research is deemed to be warranted. Given the heterogeneous nature of the autistic population, and as other authors have previously suggested (e.g. Gray, 1994b; Vivian et al, 2012; Lewandowski et al, 2014), it appears feasible that a number of factors, such as context and participant characteristics, may interact with CSC topic and content to mediate the impact of the intervention.

6.2 Unique Contribution

The present research has made a contribution to the currently very limited evidence base for the effectiveness of CSCs. Much of the existing research is non-experimental in nature, therefore the adoption of a SCED methodology has arguably added a more robust evaluation to the existing literature. This research is the first experimental study conducted in the UK context with primary-aged pupils with ASC in mainstream schools. It has also been the first study to explicitly explore the relative impacts of intervention frequency on behaviour change.

In addition to offering a unique contribution to the existing body of research relating to CSCs, this research illustrates the opportunities that SCEDs offer

practitioners working in applied settings such as schools, and their potential to make an important contribution to the evidence-based practice agenda.

References

Ahmed-Husain, S., & Dunsmuir, S. (2014). An evaluation of the effectiveness of Comic Strip Conversations in promoting the inclusion of young people with autism spectrum disorder in secondary schools. *International Journal of Developmental Disabilities*, 60(2), 89-108.

Ali, S., & Frederickson, N. (2006). Investigating the evidence base of social stories. *Educational Psychology in Practice*, 22(4), 355-377.

All Party Parliamentary Group on Autism (2012). *The Right Start: Reforming the System for Children with Autism.* National Autistic Society.

Altman, D. G. (1999). *Practical Statistics for Medical Research*. New York, NY: Chapman & Hall/CRC Press.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th Ed.).* Washington, DC: American Psychiatric Association.

Anagnostou, E., Jones, N., Huerta, M., Halladay, A. K., Wang, P., Scahill, L., Horrigan, J.P., Kasari, C., Lord, C., Choi, D., Sullivan, K., & Dawson, G. (2015). Measuring social communication behaviors as a treatment endpoint in individuals with autism spectrum disorder. *Autism*, 19(5), 622-636.

APA (no date). Retrieved from http://www.psychiatry.org/autism

Astington, J. W., & Baird, J. A. (Eds) (2005). *Why Language Matters for Theory of Mind.* New York: Oxford University Press.

Autism Education Trust (2008). *Educational Provision for Children and Young People on the Autism Spectrum Living in England: A Review of Current Practice, Issues and Challenges*. AET.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a

population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, 368(9531), 210-215.

Baranek, G. T. (2002). Efficacy of sensory and motor interventions for children with autism. *Journal of Autism and Developmental Disorders*, 32(5), 397-422.

Barker, C., Pistrang, N., & Elliott, R. (2002). *Research Methods in Clinical Psychology*. London: Wiley.

Barlow, D.H., Nock, M.K., & Hersen, M. (2009). *Single Case Experimental Designs: Strategies for Studying Behaviour Change.* Pearson Education, Inc.

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, 30(2), 285-297.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition*, 21(1), 37-46.

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry*, 194(6), 500-509.

Bellini, S., Peters, J. K., Benner, L., & Hopf, A. (2007). A meta-analysis of school-based social skills interventions for children with autism spectrum disorders. *Remedial and Special Education*, 28(3), 153-162.

British Psychological Society (2010). *Code of Human Research Ethics*. Leicester: BPS.

Brossart, D. F., Parker, R. I., Olson, E. A., & Mahadevan, L. (2006). The relationship between visual analysis and five statistical analyses in a simple AB single-case research design. *Behavior Modification*, 30(5), 531-563.

Brossart, D. F., Vannest, K. J., Davis, J. L., & Patience, M. A. (2014). Incorporating nonoverlap indices with visual analysis for quantifying intervention effectiveness in single-case experimental designs. *Neuropsychological Rehabilitation*, 24(3), 464-491.

Burden, R. (2015). All that glitters is not gold: why randomised controlled trials (RCTs) are of limited value in judging the effectiveness of literacy interventions. *Educational and Child Psychology,* 32, 11–20.

Cohen J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd edition). Hillsdale, N.J: Lawrence Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education*. Routledge.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand McNally.

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23-45.

Department for Education (2012). *Profiles of Need and Provision for Children with Language Impairments and Autism Spectrum Disorders in Mainstream Schools: A Prospective Study.* London: Department for Education.

Diehl, S. F. (2003). The SLP's role in collaborative assessment and intervention for children with ASC. *Topics in Language Disorders*, 23(2), 95-115.

Dixon, M. R., Jackson, J. W., Small, S. L., Horner-King, M. J., Lik, N. M. K., Garcia, Y., & Rosales, R. (2009). Creating single-subject design graphs

in Microsoft Excel™ 2007. *Journal of Applied Behavior Analysis*, 42(2), 277-293.

Dunsmuir, S., Brown, E., Iyadurai, S., & Monsen, J. (2009). Evidence-based practice and evaluation: from insight to impact. *Educational Psychology in Practice*, 25(1), 53-70.

EPPI Centre (no date). *What is a Systematic Review?* Retrieved from: http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=67

Ferguson, C. J. (2009). An effect size primer: a guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), 532 – 538.

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

Fox, M. (2011). Practice-based evidence - overcoming insecure attachments. *Educational Psychology in Practice,* 27(4), 325–335.

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

Frederickson, N. & Cline, T. (2009). *Special Educational Needs, Inclusion and Diversity*. McGraw-Hill Education (UK).

Frederickson, N., Miller, A. (2008). What do educational psychologists do? In Frederickson, N., Miller, A., & Cline, T. (Eds). *Educational Psychology.* Hodder Education.

Friman, P. C. (2009). Behavior Assessment. In D. H. Barlow, M. K. Nock, & M. Hersen (Eds.) *Single Case Experimental Designs: Strategies for Studying Behaviour Change* (3rd ed., pp. 99–134). Boston: Pearson Education.

Frith, U. (1989, 2003). *Autism: Explaining the Enigma.* Blackwell Publishing.

Frith, U., Happé, F., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development*, 3(2), 108-124.

Glaeser, B. C., Pierson, M. R., & Fritschmann, N. (2003). Comic Strip Conversations: A Positive Behavioral Support Strategy. *Teaching Exceptional Children*, 36(2), 14-19.

Gordon, A. C., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology*, 68(1), 70-83.

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

Gray, C. (1994a). *The New Social Stories Book*. Arlington: Future Horizons Inc.

Gray, C. (1994b). *Comic Strip Conversations*. Arlington: Future Horizons.

Gray, C. A. (1998). Social stories and comic strip conversations with students with Asperger syndrome and high-functioning autism. In *Asperger Syndrome or High-Functioning Autism?* (pp. 167-198). Springer US.

Gray, C. (2000). *The New Social Story Book*. Future Horizons.

Gray, C. (2004). Social stories 10.0: The new defining criteria and guidelines. *Jenison Autism Journal*, 15(4), 2-21.

Gray, C. A., & Garand, J. D. (1993). Social stories: Improving responses of students with autism with accurate social information. *Focus on Autistic Behavior*, 8(1), 1-10.

Grix, J. (2002). Introducing students to the generic terminology of social research. *Politics*, 22(3), 175-186.

Guba E., & Lincoln Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluence. In Denzin N. K., Lincoln Y. S. (Eds.). *The*

Sage handbook of qualitative research (3rd ed.) Thousand Oaks, CA: Sage.

Hannah, L. (2001). *Teaching Young Children with Autistic Spectrum Disorders to Learn: A Practical Guide for Parents and Staff in Mainstream Schools and Nurseries.* National Autistic Society.

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, 66(3), 843-855.

Harrington, R. (2001). Commentary: Evidence-based child and adolescent mental health services. *Child Psychology and Psychiatry Review*, 6(2), 65-65.

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

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *The British Journal of Developmental Psychology*, 16(2), 233-253.

Hutchins, T. L., Bonazinga, L. A., Prelock, P. A., & Taylor, R. S. (2008). Beyond false beliefs: The development and psychometric evaluation of the Perceptions of Children's Theory of Mind Measure-Experimental Version (PCToMM-E). *Journal of Autism and Developmental Disorders*, 38(1), 143-155.

Hutchins, T. L., & Prelock, P. A. (2006). Using social stories and comic strip conversations to promote socially valid outcomes for children with autism. *Seminars in Speech and Language*, 27(1), 47-59.

Hutchins, T. L., & Prelock, P. A. (2008). Supporting theory of mind development: Considerations and recommendations for professionals providing services to individuals with autism spectrum disorder. *Topics in Language Disorders*, 28(4), 340-364.

Hutchins, T. L., & Prelock, P. A. (2012). Parents' perceptions of their children's social behavior: The social validity of Social Stories™ and comic strip conversations. *Journal of Positive Behavior Interventions*, 15(3), 156-168.

Hutchins, T. L., Prelock, P. A., & Bonazinga, L. (2012). Psychometric evaluation of the Theory of Mind Inventory (ToMI): A study of typically developing children and children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 42(3), 327-341.

Jones, G. (2002). *Educational Provision for Children with Autism and Asperger Syndrome: Meeting Their Needs*. Routledge.

Jones, G. (2006). Department for Education and Skills/Department of Health: Good practice guidance on the education of children with autistic spectrum disorder. *Child: Care, Health and Development*, 32(5), 543-552.

Jordan, R. R. (2007). A rose by another name? *Communication*, 12-15.

Karkhaneh, M., Clark, B., Ospina, M. B., Seida, J. C., Smith, V., & Hartling, L. (2010). Social Stories™ to improve social skills in children with autism spectrum disorder: a systematic review. *Autism*, 14(6), 641-662.

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

Kerr, S., & Durkin, K. (2004). Understanding of thought bubbles as mental representations in children with autism: Implications for theory of mind. *Journal of Autism and Developmental Disorders*, 34(6), 637-648.

Kokina, A., & Kern, L. (2010). Social Story Interventions for Students with Autism Spectrum Disorders: A Meta-Analysis. *Journal of Autism and Developmental Disorders,* 40(7), 812-826.

Kratochwill, T.R. (1992) Single-case research design and analysis: an overview. In Kratochwill, TR., and Levin, J.R. (Eds). *Single-Case Research Design and Analysis.* Lawrence Erlbaum: Hillsdale, NJ.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation.* What Works Clearing House.

Kratochwill, T. R., Hitchcock, J. H., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2013). Single-case intervention research design standards. *Remedial and Special Education*, 34(1), 26-38.

Leslie, A. M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology*, 6(4), 315-324.

Lewandowski, J. F., Hutchins, T. L., Prelock, P. A., & Murray-Close, D. (2014). Examining the Benefit of Including a Sibling in Story-Based Interventions with a Child With Asperger Syndrome. *Contemporary Issues in Communication Science & Disorders*, 41, 179-195.

Ma, H. H. (2006). An alternative method for quantitative synthesis of single-subject researches percentage of data points exceeding the median. *Behavior Modification*, 30(5), 598-617.

Manolov, R., & Solanas, A. (2008). Comparing N=1 effect size indices in presence of autocorrelation. *Behavior Modification*, 32(6), 860-875.

Manolov, R., Solanas, A., Sierra, V., & Evans, J. J. (2011). Choosing among techniques for quantifying single-case intervention effectiveness. *Behavior Therapy*, 42(3), 533-545.

Medical Research Council (2001). *Review of Autism Research: Epidemiology and Causes.* London: Medical Research Council.

Mertens, D. M. (2010). *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*. Sage Publications Ltd.

Mesibov, G. B., & Howley, M. (2003). *Accessing the Curriculum for Pupils with Autistic Spectrum Disorders: Using the TEACCH Programme to Help Inclusion.* David Fulton Publishers.

National Autistic Society (no date). Retrieved from http://www.autism.org.uk/about-autism/myths-facts-and-statistics/statistics-how-many-people-have-autism-spectrum-disorders.aspx

Neef, N. (2009). Advances in single subject design. In Gutkin, T.B., & Reynolds, C.R. (Eds) (2009). *The Handbook of School Psychology.* John Wiley & Sons.

Nock, M. K., Michel, B. D., & Photos, V. I. (2007). Single-case research designs. In D. McKay (Ed.). *Handbook of Research Methods in Abnormal and Clinical Psychology.* Thousand Oaks, CA: Sage.

Odom, S. L., Brown, W. H., Frey, T., Karasu, N., Smith-Canter, L. L., & Strain, P. S. (2003). Evidence-based practices for young children with autism contributions for single-subject design research. *Focus on Autism and Other Developmental Disabilities*, 18(3), 166-175.

Onghena, P., & Edgington, E. S. (2005). Customization of pain treatments: Single-case design and analysis. *The Clinical Journal of Pain*, 21(1), 56-68.

Ozonoff, S. (1997). Components of executive function in autism and other disorders. In Russel, J. (Ed). *Autism as an Executive Disorder*. Oxford: Oxford University Press.

Ozonoff, S., & Miller, J. N. (1995). Teaching theory of mind: A new approach to social skills training for individuals with autism. *Journal of Autism and Developmental Disorders,* 25(4), 415-433.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND) an alternative to PND. *The Journal of Special Education*, 40(4), 194-204.

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The improvement rate difference for single-case research. *Exceptional Children*, 75(2), 135-150.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, 42(2), 284-299.

Parsons, S., & Mitchell, P. (1999). What children with autism understand about thoughts and thought bubbles. *Autism*, 3(1), 17-38.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 37(1), 51-87.

Perrin, T. (1998) Single-system methodology: a way forward in dementia care? *British Journal of Occupational Therapy*, 61 (10), 448-452.

Petticrew, M., & Roberts, H. (2008). *Systematic Reviews in the Social Sciences: A practical Guide*. John Wiley & Sons.

Pierson, M. R., & Glaeser, B. C. (2005). Extension of research on social skills training using comic strip conversations to students without autism. *Education and Training in Developmental Disabilities*, 40(3), 279-284.

Pierson, M. R., & Glaeser, B. C. (2007). Using comic strip conversations to increase social satisfaction and decrease loneliness in students with autism spectrum disorder. *Education and Training in Developmental Disabilities*, 42(4), 460-466.

Quill, K. A. (1997). Instructional considerations for young children with autism: The rationale for visually cued instruction. *Journal of Autism and Developmental Disorders*, 27(6), 697-714.

Rao, P. A., Beidel, D. C., & Murray, M. J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: A review and recommendations. *Journal of Autism and Developmental Disorders*, 38(2), 353-361.

Reichow, B., Steiner, A. M., & Volkmar, F. (2013). Cochrane review: social skills groups for people aged 6 to 21 with autism spectrum disorders (ASC). *Evidence-Based Child Health: A Cochrane Review Journal*, 8(2), 266-315.

Reynhout, G., & Carter, M. (2006). Social Stories for children with disabilities. *Journal of Autism and Developmental Disorders,* 36(4), 445-469.

Reynhout, G. & Carter, M. (2011). Evaluation of the efficacy of Social Stories using three single subject metrics. *Research in Autism Spectrum Disorders,* 5(2), 885-900.

Robinson, S. (2008). Using a strategy of 'structured conversation' to enhance the quality of tutorial time. *Journal of Further and Higher Education*, 32(1), 59-69.

Robson, C. (2011). *Real World Research*. Wiley.

Rogers, S. J. (2000). Interventions that facilitate socialization in children with autism. *Journal of Autism and Developmental Disorders*, 30(5), 399-409.

Rogers, M. F., & Myles, B. S. (2001). Using social stories and comic strip conversations to interpret social situations for an adolescent with Asperger syndrome. *Intervention in School and Clinic*, 36(5), 310.

Ross, S. G., & Begeny, J. C. (2014). Single-case effect size calculation: comparing regression and non-parametric approaches across previously published reading intervention data sets. *Journal of School Psychology*, 52(4), 419-431.

Roth, A., & Fonagy, P. (1996). *What Works for Whom? A Critical Review of Psychotherapy Research.* New York: The Guilford Press.

Scattone, D., Wilczynski, S. M., Edwards, R. P. and Rabian, B. (2002). Decreasing disruptive behaviors of children with autism using Social Stories. *Journal of Autism and Developmental Disorders*, 32(6), 535-543.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single-subject research methodology and validation. *Remedial and Special Education*, 8(2), 24-33.

Smith, C. (2001). Using social stories to enhance behaviour in children with autistic spectrum difficulties. *Educational Psychology in Practice*, 17(4), 337-345.

Stoiber, K. C., & Waas, G. A. (2002). A contextual and methodological perspective on the evidence-based movement within school psychology in the United States. *Educational and Child Psychology*, 19 (3), 7-21.

Thiemann, K. S., & Goldstein, H. (2001). Social Stories, written text cues and video feedback: Effects on social communication of children with autism. *Journal of Applied Behavior Analysis*, 34(4), 425-446.

Todman, J. B., & Dugard, P. (2001). *Single-Case and Small-n Experimental Designs: A Practical Guide to Randomization Tests*. Psychology Press.

Tse, J., Strulovitch, J., Tagalakis, V., Meng, L., & Fombonne, E. (2007). Social skills training for adolescents with Asperger syndrome and high-functioning autism. *Journal of Autism and Developmental Disorders*, 37(10), 1960-1968.

University of Nottingham (2013). *Code of Research Conduct and Research Ethics.* University of Nottingham.

Vannest, K. J., & Ninci, J. (2015). Evaluating Intervention Effects in Single-Case Research Designs. *Journal of Counseling and Development*, *93*(4), 403-411.

Vivian, L., Hutchins, T. L., & Prelock, P. (2012). A family-centered approach for training parents to use comic strip conversations with their child with

autism. *Contemporary Issues in Communication Science and Disorders*, 39, 30-42.

Volkmar, F. R. (2011). Understanding the social brain in autism. *Developmental Psychobiology*, 53(5), 428-434.

Wellman, H. M., Baron-Cohen, S., Caswell, R., Gomez, J. C., Swettenham, J., Toye, E., & Lagattuta, K. (2002). Thought-bubbles help children with autism acquire an alternative to a theory of mind. *Autism*, 6(4), 343-363.

White, S. W., Keonig, K., & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: A review of the intervention research. *Journal of Autism and Developmental Disorders*, 37(10), 1858-1868.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, 9(1), 11-29.

Wing, L., & Potter, D. (2002). The epidemiology of autistic spectrum disorders: is the prevalence rising? *Mental Retardation and Developmental Disabilities Research Reviews*, 8(3), 151-161.

Secondary References

Asperger, H. (1944). Die autistischen psychopathen im kindesalter. In Frith, U. (2003). *Autism: Explaining the Enigma.* Blackwell Publishing.

Kanner, L. (1943). Autistic disturbances of affective contact. In Frith, U. (2003). *Autism: Explaining the Enigma.* Blackwell Publishing.

Appendices

Appendix 1: Sally Anne False Belief Task

this is Sally   this is Anne

Sally puts her ball in the basket

Sally goes away

Anne moves the ball to her box

where will Sally look for her ball?
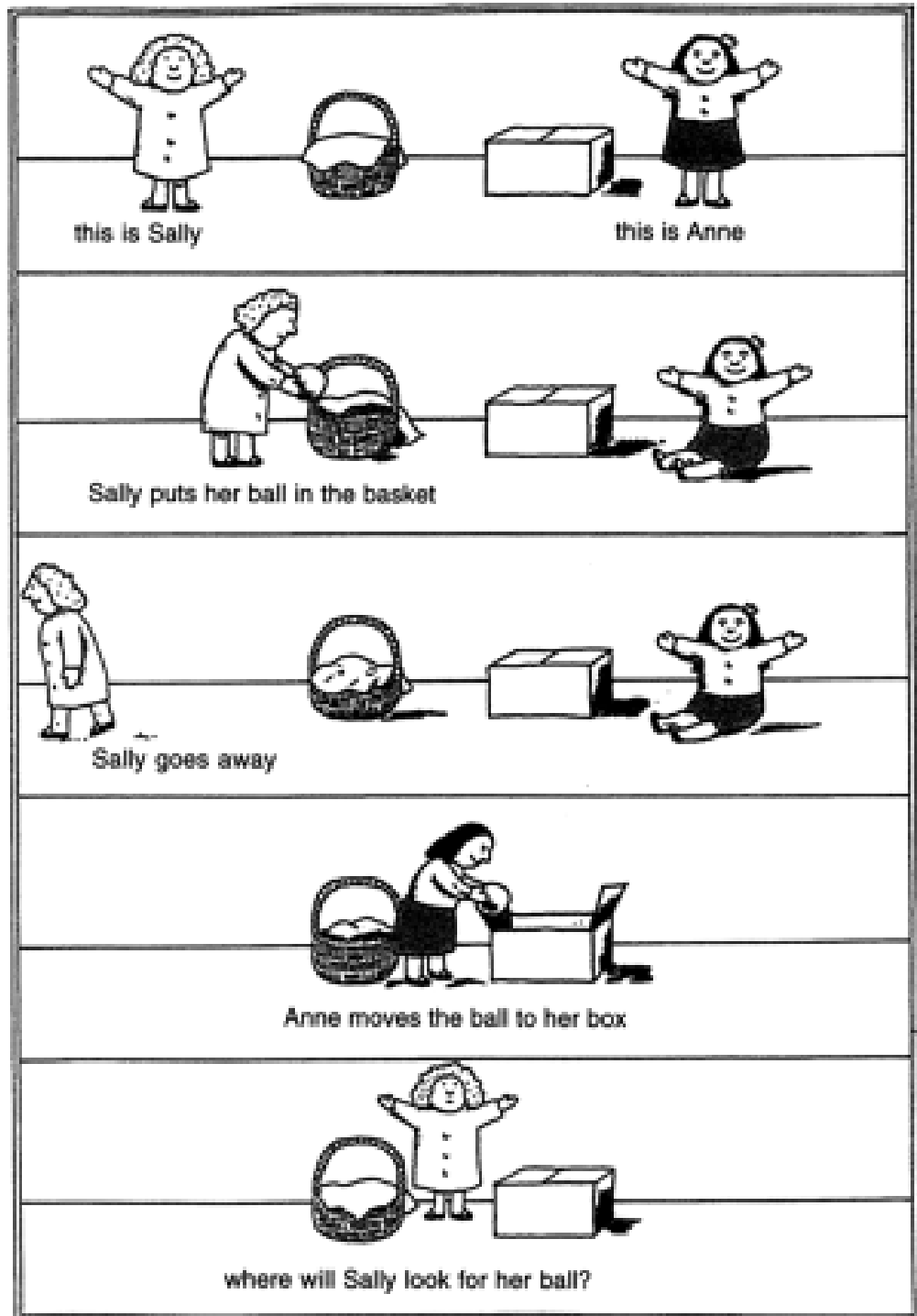
Appendix 2: Examples of CSCs (taken from the present research)

Date: 8-2-16.   Topic of conversation: Touching clay models belonging to other pupils.
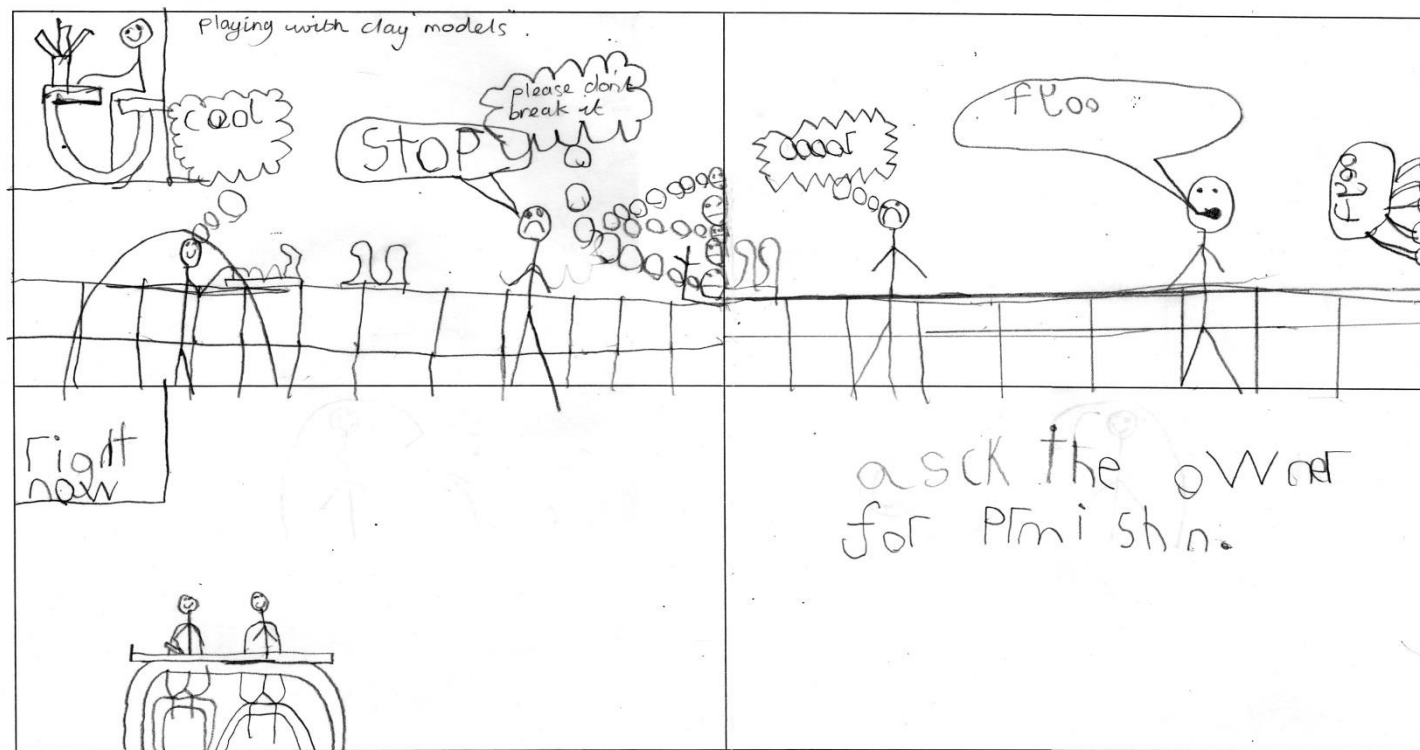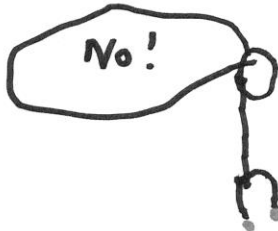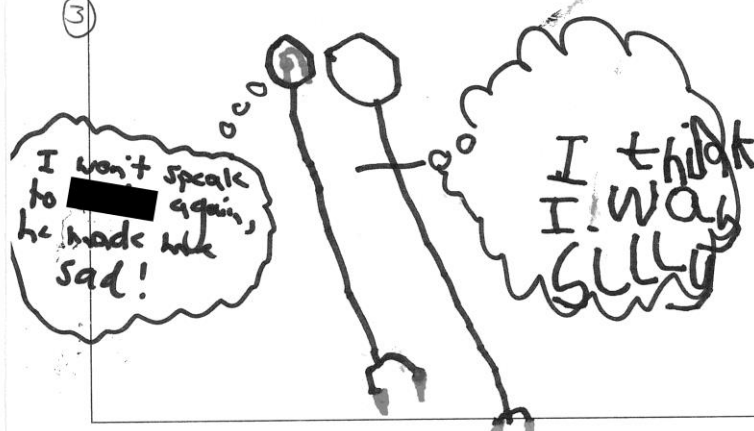
181

182

I like doing THIS iT make me feel happy.
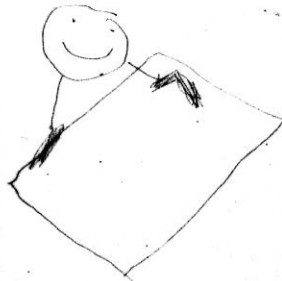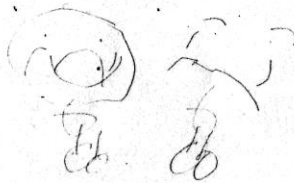
we Done like waving his finger by our face. OUR

iT makes us frightone

iT makes me feel scar too.

Appendix 3: Articles Excluded from the Systematic Literature Review

Search Result	Reason for Exclusion
Attwood (2000)	Discussion paper
Butterly (2012)	Book review
De La Iglesia & Olivar (2008)	Spanish Language / Review paper
Deshpande, Libero, Sreenivasan, Deshpande & Kana (2013)	Not relevant to review question
Duverger, Da Fonseca, Bailly & Deruelle (2007)	Not relevant to review question
Herbet, Lafargue, Bonnetblanc, Moritz-Gasser, de Champfleur & Duffau (2014)	Not relevant to review question
Hutchins & Prelock (2006)	Intervention included Social Stories
Hutchins & Prelock (2008)	Intervention included Social Stories
Hutchins & Prelock (2012)	Intervention included Social Stories
Kana, Libero, Hu, Deshpande & Colburn (2014)	Not relevant to review question
Lequia (2011)	Book review
McConkey (2006)	Book review
Philpott, Rinehart, Gray, Howlin & Cornish (2013)	Not relevant to review question
Pierson & Glaeser (2005)	Participants did not have ASC diagnoses
Sivaratnam, Cornish, Gray, Howlin & Rinehart (2012)	Not relevant to review question
Sivaratnam, Cornish & Gray (2012)	Not relevant to review question
Strick (2012)	Book review
VanBergeijk (2010)	Book review
Whittingham, Sofronoff, Sheffield & Sanders (2009)	No outcome measures relating to CSCs
Zucker, Perras, Gartin & Fidler (2005)	Discussion paper

Appendix 4: Systematic Review - Key Facts Table

Study	Design	Participants & Setting	Procedural information	Target behaviours	Measure(s)	Results
Rogers & Myles (2001)	Case study	14 year old male student with a diagnosis of Asperger syndrome who attended a combination of general and special educational contexts in the U.S.	A social story intervention was first introduced by itself before a CSC intervention was introduced. CSCs were completed daily with a teacher. The article suggests that the teacher completed the drawing, not the pupil. The duration period of the intervention unclear.	Difficulties in 'settling' after lunch periods and arriving late to class as a result.	Event sample of the number of verbal redirections needed following a lunch break Timing of the number of minutes late to P.E class	Number of redirections decreased following implementation of social stories, then further decreased following implementation of CSC. Student reported enjoying using CSCs and began to request their use at home and school.
Glaeser, Pierson & Fritschmann (2003)	Descriptive case studies	2nd grade female student with 'mild autism' who attended a general education setting in the U.S	Female student used CSCs at least once a day with a staff member, even if no conflict had occurred	Conflicts with peers and adults	Unstructured observations / anecdotal data collected by staff	Participant showed 'great progress in dealing with conflicts with

		There was also a case description for a second student who did not have a diagnosis of ASC - these results are not reported.	(in which case any sort of interaction was drawn and written about). Overall duration of use is unclear.			peers and adults' over the first few months of use.
Pierson & Glaeser (2007)	Case studies	3 students aged 6,7 and 8 years old with high-functioning autism Special educational setting (in U.S)	CSCs created alongside a teacher or teaching paraprofessional after any negative social experience occurred in the classroom or playground. Appropriate social behaviours were reinforced throughout the day by staff reminding them of a previous CSC's solution. 6 week intervention period	Participant 1: Appropriate use of hands and feet when playing games on the playground Participant 2: Increase use of social greeting with eye contact and appropriate voice volume Participant 3: Accepting responsibility for inappropriate	Observations carried out by teacher or paraprofessional	Data indicated 'significant changes' in target behaviours for all participants (between 50% - 75% improvement from baseline) Anecdotal records for each participant were analysed for decreased levels of loneliness - this indicated visible signs of social

				actions and apologising to peers It was determined that these target behaviours had an impact on social competence and, as a result, feelings of loneliness and satisfaction with relationships.		satisfaction, including fewer loneliness verbalisations, increased talkativeness with peers, more smiles, and a greater desire to participate with peers in the classroom and on the playground.
Robinson (2008)	Case study	2 students with ASC diagnoses Study carried out in a Higher Education context in the UK	CSCs were used by each student and their tutor to structure tutorials	Communicative intent and interaction relevant to the content of tutorials	Direct observation by researcher	Both students demonstrated increased levels of communicative intent and interaction relevant to the content of

						tutorials during the 'CSC tutorial' compared to a typical tutorial.
Vivian, Hutchins & Prelock (2012)	Case study	8 year old female and her parents Family context in the U.S	Parents implemented CSC intervention over a 6 week period (included both positive and negative situations) A total of 20 CSCs were created.	CSCs focused on a wide variety of behaviours, e.g. trouble at bedtime, talking back to teachers, hitting friends, responding to requests even when you do not want to do something, getting 'time-out' in school.	Theory of Mind Inventory, Post-intervention interview with parents	Participant scored at the 50th percentile on the Theory of Mind Inventory at post-intervention, compared to the 17th percentile pre-intervention Parents considered CSCs to be a feasible intervention which 'definitely helped' facilitate more appropriate behaviours.
Ahmed-Husain & Dunsmuir	ABA multiple-baseline	8 students (3 female, 5 male) aged 11-14 years with diagnoses of ASC	Each participant created two CSCs (alongside staff	Range of behaviours targeted including	Interval sampling observations conducted jointly	Visual analysis, percentage of data points exceeding

| (2014) | across behaviours single case design | or Aspergers who attended mainstream secondary schools in the UK. | members) focusing on different target behaviours. One CSC contained a visual action plan, the other an auditory action plan.

CSCs reviewed 3-5 x per week. Each CSC was implemented for between 4-6 weeks. | making eye contact, fiddling with objects in lessons, banging and tapping, making inappropriate comments to others, shutting eyes and putting head on desk in lessons, asking for help too often, initiating conversations | by researcher and teacher. | the mean and Tau-U analyses indicated the intervention was found to be moderately to highly effective at reducing/improving target behaviours in 7 out of 8 participants.

4 out of 8 participants responded well to both CSCs; 3 out of 8 responded well to one of the CSCs; for 1 participant neither CSCs were effective

Participants verbal and visual skills matched the |

						type of action plan used for the more successful CSC
Lewandowski, Hutchins, Prelock & Murray-Close (2014)	ABACA single-case design which also included some qualitative, ethnographic methods	5 year old male with a diagnosis of Asperger syndrome and his typically developing younger brother (the results for the latter are not reported) The intervention took place in the family home in the U.S	A total of 12 CSCs in each intervention phase were delivered one to two times per week. A few affirmative CSCs were also included. CSCs were created with a researcher working alongside the child on a topic suggested by their mother. Length of phases: A = 34 days B = 89 days A = 38 day C = 75 days A = 77 days	Sibling conflict	Daily diaries completed by parent which included a 10-point Likert-type rating scale Theory of Mind Inventory	Tau-U analysis indicated no treatment effect from the A to B phase but a statistically significant difference between the second A phase and the C phase for the participant with Asperger syndrome. This effect was maintained over the final withdrawal phase.

Appendix 5: Letter to Headteachers

Dear Headteacher,

My name is Joanne Page and I am a Trainee Educational Psychologist currently working for -------- Educational Psychology Service as part of my Doctorate in Applied Educational Psychology that I am completing through the University of Nottingham. I am writing to inform you about a research study that I am going to be undertaking and to invite your school to participate in it.

The aim of the research is to evaluate the effectiveness of Comic Strip Conversations (CSCs) in addressing the target social behaviours of pupils with a diagnosis of autism. CSCs use visual representations, such as symbols, stick figures and colours, to review situations taken directly from an individual's life and to discuss alternatives to behaviour which had been unbeneficial in that situation. They encourage children to consider other people's thoughts and feelings in a situation as well as their own. The overall aim is to help teach social skills and improve social understanding. At present, the evidence base for the effectiveness of CSCs is very small but, of the research that does exist, positive outcomes have been indicated. My research aims to add to the existing evidence base and investigate how effective CSCs are in addressing specific, target behaviours. In addition, I am also interested in investigating how the effects of the intervention may vary depending on how frequently CSCs are created.

My aim is to work alongside school staff to identify children in school with a diagnosis of autism who it is thought might benefit from such an intervention. I am looking for primary-aged participants who are frequently displaying behaviours that school are concerned about (e.g. disagreements with others, not following instructions). Alternatively, it may be that the absence of certain behaviours is more of a concern (e.g. not greeting others, not initiating communication).

The process would involve:
- an initial meeting with the appropriate school staff and parent(s)/carer(s) to provide more detailed information about the project and to gather information regarding the pupil's specific behaviours.
- observation of the pupil carried out by myself to collect further information about their behaviours.
- either myself or a member of staff conducting regular observations to record the occurrence of the target behaviour over a period of

approximately two to three weeks before the CSC intervention is implemented, and then continuing this over a period of approximately three to four weeks whilst the intervention is being implemented. The precise details of this would be decided once more information about the target behaviours has been identified.

- a member of school staff completing one CSC with the pupil for one of their target behaviours, with daily reviews, and weekly CSCs for the second target behaviours. It should take no longer than thirty minutes to complete a CSC, and reviews will only last approximately five minutes.

All of the work will be carried out professionally in line with the ethical guidelines of the British Psychological Society, which includes ensuring the confidentiality of data and the right to withdraw from the study at any time. Parents/Carers will be fully informed about the nature of the research and will be invited to give consent for their child to participate. The children involved will also be provided with information and their consent will be gained. On completion of the research, I will provide written feedback summarising the study and its findings and I can provide additional feedback in person as well if wanted.

I hope this letter has been useful in providing you with an initial idea of what the research would entail and the potential benefits that such an intervention could have for your staff and the children and families involved, in addition to assisting me in my training.

If you feel that you have any pupils for whom you think my research may benefit, then please do contact me at ------------- or ------------- and we can arrange a meeting to discuss it in more detail. You are welcome to contact me with any questions you may have. The contact details of my research supervisor, Dr. Sarah Atkinson, are also supplied below.

Thank you for your time.

Yours sincerely,

Joanne Page (Trainee Educational Psychologist)

<contact details>

Dr. Sarah Atkinson (Academic and Professional Tutor)

Appendix 6: Information letter for parents and parental consent form

Dear Parent/Guardian,

My name is Joanne Page and I am a trainee educational psychologist currently working for ---------- Educational Psychology Service as part of the Doctorate in Applied Educational Psychology that I am completing through the University of Nottingham. I am writing to inform you about a research study that I am going to be undertaking.

The aim of my research is to evaluate the effectiveness of Comic Strip Conversations (CSCs) in addressing the target social behaviours of pupils with a diagnosis of Autistic Spectrum Condition (ASC). CSCs use visual representations, such as symbols, stick figures and colours, to review situations taken directly from an individual's life and to discuss alternatives to behaviour which had proven to be unbeneficial in that situation. They encourage children to consider other people's thoughts and feelings in a situation as well as their own. The overall aim is to help teach social skills and improve social understanding.

The study would involve me working alongside yourself and school staff to identify target behaviours that may benefit from being addressed through a CSC intervention. The process would involve:
- an initial meeting with yourself and school staff to gather information regarding the specific behaviours that your child presents with in school that may benefit from targeted intervention.
- observation of your child carried out by myself to collect further information about their behaviours.
- either myself or a member of staff conducting observations to record the occurrence of the target behaviour over a period of approximately two to three weeks before the CSC intervention is implemented, and then continuing this over a period of approximately three to four weeks whilst the intervention is being implemented. The precise details of this can be decided once more information about the target behaviours has been identified.
- a member of school staff completing one CSC with your child for one of their target behaviours, with daily reviews, and weekly CSCs for the second target behaviour. It should take no longer than 30 minutes to complete a CSC, and reviews will only last a few minutes.

All of the work will be carried out professionally in line with the ethical guidelines of the British Psychological Society, which includes ensuring the confidentiality of the data collected and the right to withdraw from the study at

any time without needing to provide a reason. I will also seek to inform your child and gain their consent in a way that is most appropriate for them. On completion of the research, I will provide written feedback summarising the study and its findings, and would be happy to provide additional feedback in person as well.

I hope this letter has been useful in providing you with an initial idea of what the research would entail and the potential benefits that such an intervention could have for your child and the school staff working with them, in addition to assisting me in my training. Please contact me if you wish to meet or speak in person to further discuss this research.

If you are happy for your child to participate, could you please complete the enclosed consent form and return it to your child's school.

Thank you for your time.

Yours sincerely,
Joanne Page (Trainee Educational Psychologist)

<contact details>

Dr. Sarah Atkinson (Academic and Professional Tutor)

Parental Consent Form

Project title: An investigation into the effectiveness of Comic Strip Conversations for improving the target social behaviours of primary-aged pupils on the autistic   spectrum.

Researcher's name: Joanne Page

Supervisor's name: Dr Sarah Atkinson

Please read the following statements carefully:

- I have been informed of and understand the purposes of the research
- I have been given an opportunity to ask questions and to speak to or meet with the researcher.
- I understand I can withdraw my child from the research at any time without needing to provide a reason.
- I understand that any information and data gained during the study will be kept anonymous and confidential.
- I understand that I may contact the researcher or supervisor if I require further information about the research.
- I consent for my child to participate in the study as outlined to me.

Name of child: ………………………………………….

Name of parent/guardian: ……………………………………

Signature: ………………………………………………………

Date: ……………………..

Contact details:

Researcher: -----------------------

Supervisor: ----------------------

Appendix 7: Semi-structured interview proforma

Potential Participant Information Gathering Proforma

Date:

Present:

Name	
D.O.B	
Setting	
Date of diagnosis	
Any additional diagnosis?	
Statemented?	
What is the level/nature of current support in school? *-Interventions/strategies?* *-How long for?* *-Impacts?*	
How would you describe their expressive & receptive language skills?	
What possible target behaviours are there? *-When/where do these occur?* *-With whom?* *-Frequency/duration?* *-Effects on learning/relationships?*	

Any additional information:

Appendix 8: Participant information letter and consent form

I would like to invite you to take part in a study about Comic Strip Conversations.

This letter will give you important information about what will happen if you take part. Your teacher(s) will also talk to you about it to make sure that you understand everything in this letter.

If you take part, then you will be creating Comic Strip Conversations with an adult in school. This will involve drawing and talking about situations that have happened to you in school. These might be situations where things did not go very well for you. You will be thinking about what you did and how you felt in these situations. You will also think about how other people may have been feeling. You get to think about what you could do differently if something similar happened again.

An adult will help you make your comic strips. It might take about half an hour (maybe less) to create a comic strip, but you won't be making one everyday. However, everyday an adult will sit with you and quickly remind you of the comic strips that you have already made. You will be using Comic Strip Conversations for about four to six weeks. We will be interested in finding out how you are getting on with the Comic Strip Conversations, so information about how you are behaving in school will be being collected.

You do not have to take part if you do not want to. You can leave the study at any time without having to say why. Any information that is collected about you will be kept confidential - this means that no one will be able to tell that it is about you.

You can ask any questions that you have at any time. There will be someone in school that you can speak to about it (someone will let you know who) and you can also talk to your parents or to me.

Thank you for your time.

Joanne Page (Trainee Educational Psychologist)

<div align="center">Participant Consent Form</div>

Title of Project: An investigation into the effectiveness of Comic Strip Conversations for improving the target social behaviours of primary-aged pupils on the autistic spectrum

<div align="center">

Ethics Approval Number: 635

Researcher(s): Joanne Page, *(*contact details)

Supervisor: Sarah Atkinson, (contact details)

</div>

Please read carefully and answer the questions below.

- Do you understand the information given in the Participant Information Sheet?
  YES/NO

- Have you been able ask questions about the study if you wanted to?
  YES/NO

- Are you happy with the answers you got to your questions?
  YES/NO

- Do you understand that you can leave the study at any time without giving a reason?
  YES/NO

- Are you happy for the information collected in this study to be shared with other people? Your name would not be used so people would not know that the information was about you.
  YES/NO

- Do you agree to take part in the study?
  YES/NO

"The study has been explained to me in a way that I understand, and I agree to take part. I understand that I can leave the study at any time."

Signature of the Participant: Date:

Name (in block capitals):

The study has been explained to the above participant and he/she has agreed to take part.

Signature of researcher: Date:

Appendix 9: Staff information sheet and consent form

Title of Project: An investigation into the effectiveness of Comic Strip Conversations for improving the target social behaviours of primary-aged pupils on the autistic spectrum

Researcher: Joanne Page <contact details>

Supervisor: Dr Sarah Atkinson <contact details>

Outlined below are the responsibilities of the school and the researcher the with regards to this research, and the potential benefits for the school of taking part in the research. This information is supplementary to that which can be found in the 'School Information Sheet'. Please read the information carefully before completing the attached consent form.

Responsibilities of the researcher:

- To train the staff member in the use of Comic Strip Conversations
- To train staff in the use of the observational technique to be used to collect data
- To be involved in the data collection process through conducting some joint observations of the participant (these occasions TBC)
- To keep in regular contact with the school to offer ongoing support and advice over the course of the research
- To analyse the data collected and feedback the research outcomes to all involved
-

Responsibilities of school staff:

- To implement the Comic Strip Conversation intervention with the frequency specified at each phase of the research (a one-off CSC for one target behaviour with daily reviews, and weekly CSCs for the second target behaviour).
- *To undertake regular event-sample observations of the participants for the purpose of data collection (the extent of this will be dependent on the nature of the participant's behaviours - once more information is known about the participants, this point can be personalised for each individual case)*
- Throughout the course of the research, to monitor the assent and well-being of the participants, and to contact the researcher should concerns arise.

Potential benefits to the school for taking part:

- The opportunity to take part in a piece of research which aims to inform the evidence base for the effectiveness of CSCs.
- The staff member(s) involved will be equipped with the knowledge and skills to be able to implement a practical and inexpensive intervention, which can be applied to other pupils in the school if wished.
- The opportunity to gain detailed monitoring data about the impact of the intervention which can inform future planning for this pupil, and potentially others.

Staff consent form

Title of Project: An investigation into the effectiveness of Comic Strip Conversations for improving the target social behaviours of primary-aged pupils with autism.

Researcher: Joanne Page <contact details>

Supervisor: Sarah Atkinson <contact details>

Please read carefully and answer the questions below.

- Have you read and understood the Staff Information Sheet?
  YES/NO

- Have you had the opportunity to ask questions about the study?
  YES/NO

- Have all your questions been answered satisfactorily?
  YES/NO

- Do you understand that you are free to withdraw from the study (at any time and without giving a reason)?
  YES/NO

- I give permission for the data from this study to be shared with other researchers provided that their anonymity is completely protected.
  YES/NO

- Do you agree to take part in the study?
  YES/NO

"This study has been explained to me to my satisfaction, and I agree to take part. I understand that I am free to withdraw at any time."

Signature: Date:

Name (in block capitals):

I have explained the study to the above staff member and he/she has agreed to take part.

Signature of researcher: Date:

Appendix 10: Example Observation Schedules

Jack - Interval Sampling Observation Schedule

Observe pupil for 25 seconds, then use the following 5 seconds to record whether or not the target behaviours occurred (i.e. 25 seconds observation, 5 seconds recording, 25 seconds observation, 5 seconds recording etc...). Complete over a 30 minute time period in the classroom for target behaviour 1. Complete a further 10 minute observation on the playground for target behaviour 2.

Target Behaviour 1 (B1): Not waiting turn to speak - to include behaviours such as calling out inappropriately/interrupting a member of staff when they are taking to somebody else (not to include instances of calling out if lots of other children doing the same at that time).

Target Behaviour 2 (B2): Touching/squeezing peers - to include behaviours such as putting his arms around another person and squeezing them or putting his hands on their face/head and squeezing, or other instances of uninvited touching (not to include instances of appropriately timed contact or accidental contact).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
B1																														

	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
B1																														

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
B2																				

Daniel – Event Sample Observation Schedule

Observations to be carried out twice weekly, over a 30 minute period.

Target Behaviour 1 (B1): Asking time-related enquiries (e.g. What time is it? Is it morning? Is it afternoon?)

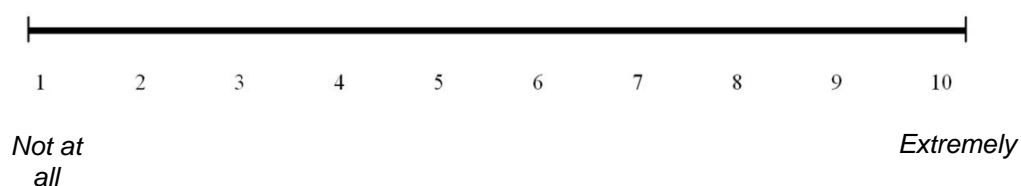**Target Behaviour 2 (B2):** Waving fingers in front of another person's face.
Not to include instances of waving his finger in front of his own face only.

	Date & time	Context (lesson/activity)	Frequency count		Total
Week No.			B1		
			B2		
			B1		
			B2		
Week No.			B1		
			B2		
			B1		
			B2		

Appendix 11: Pre- and post-intervention questionnaires

Pre-intervention questionnaire

How challenging / disruptive do you presently find the behaviour to be?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Not at all   *Extremely*

How much of an impact does the behaviour have on the pupil's everyday school life in terms of classroom learning / peer relationships

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

No impact   *Significant impact*
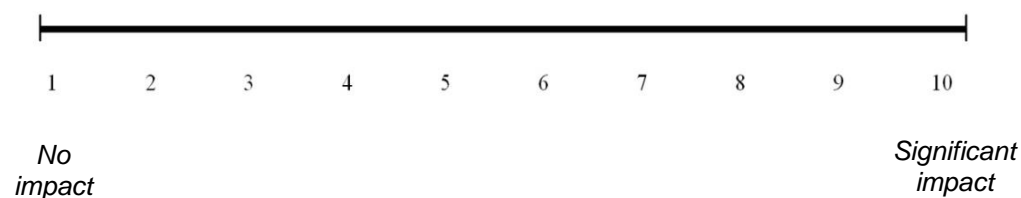
Participant Number:

Target behaviour:

Completed by (name & role):
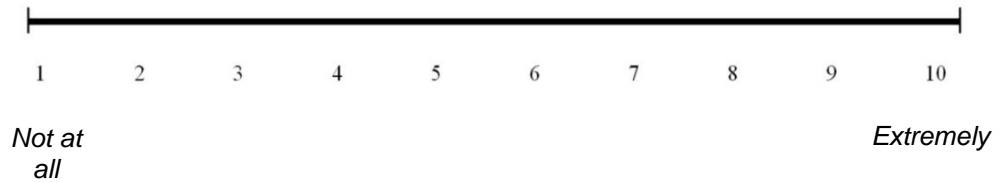
Date:

Post-intervention questionnaire

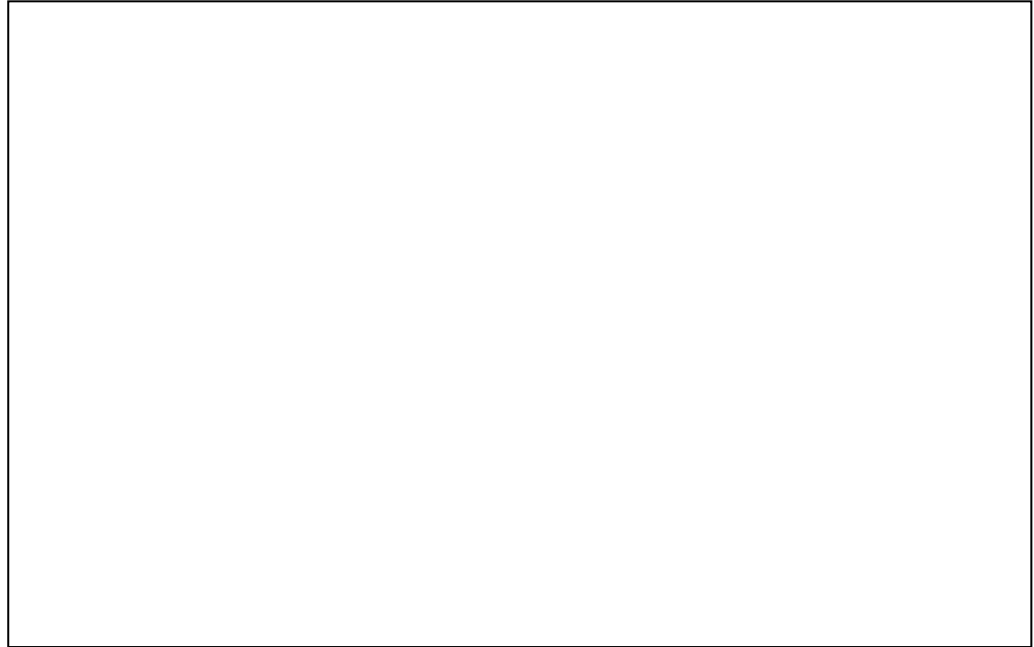How challenging / disruptive do you presently find the behaviour to be?

```
|—————————————————————————————————————————————|
  1     2     3     4     5     6     7     8     9     10
```

Not at   *Extremely*
all

How much of an impact does the behaviour have on the pupil's everyday school life in terms of classroom learning / peer relationships?

```
|—————————————————————————————————————————————|
  1     2     3     4     5     6     7     8     9     10
```

No   *Significant*
impact   *impact*

How effective have you found Comic Strip Conversations to be for addressing the target behaviour?

```
|—————————————————————————————————————————————|
  1     2     3     4     5     6     7     8     9     10
```

                                                            Highly
Not at all   *effective*
effective

Please provide any additional comments about your perceptions of the participant's response to the intervention (positive or negative) for this target behaviour:

Participant number:

Target behaviour:

Completed by (name & role):

Date:

Appendix 12: CSC training materials

Comic Strip Conversations

Training and support materials
Adapted from Gray (1994)

Comic Strip Conversations
Introduction

▸ Comic Strip Conversations (CSCs) were developed by Carol Gray (1991), based on the findings that visual supports are often useful in structuring learning for individuals with autism.

▸ A CSC is a conversation between two people which includes the use of simple drawings and other visual supports to help illustrate social skills which can prove abstract or difficult for individuals with autism to understand.

▸ In particular, CSCs aim to support children in identifying and understanding the thoughts, feelings and motivations of others.

▸ They also aim to identify possible alternative ways of behaving when similar situations arise in the future.

▸ The process may reveal misinterpretations/misunderstandings not otherwise communicated.

The Procedure

Some key points...

- The **child takes the lead** as much as is possible
- The **adult guides sensitively** without taking the lead
- The **child is encouraged to write/draw/talk the majority of the time** (although the adult also has access to pens)
- There is an **emphasis placed on what people may be thinking** as an integral part of each conversation.
- It is important to convey the message that completion of a **CSC is not a 'telling off' or a form of punishment.**
- The CSCs need a **clear structure and sequence** – comic strip boxes/numbering can be used as a guide to the order of events.

Initial conversation:
Introducing the method

- Introduce the concept of drawing whilst having a conversation – "we are going to draw while we talk today".
- Sit *next to* each other (not facing) to encourage joint attention on the task, perhaps with the adult sitting slightly back. It can be helpful for the child to have the paper clearly in front of them, to encourage the child to lead.
- Ask the child to think of something they have recently done that they would like to talk about – ideally, choose an easy topic of interest (e.g. what they did at the weekend; a trip to the park etc).
- Holding this initial conversation will be helpful in familiarising the child with the CSC method of discussing situations.

Initial conversation: Introducing the method

- Introduce the concept of drawing whilst having a conversation – "we are going to draw while we talk today".
- Sit *next to* each other (not facing) to encourage joint attention on the task, perhaps with the adult sitting slightly back. It can be helpful for the child to have the paper clearly in front of them, to encourage the child to lead.
- Ask the child to think of something they have recently done that they would like to talk about – ideally, choose an easy topic of interest (e.g. what they did at the weekend; a trip to the park etc).
- Holding this initial conversation will be helpful in familiarising the child with the CSC method of discussing situations.

Drawing about a given situation

- It can be useful to begin with some 'small talk' before introducing the topic of conversation.
- A location symbol is drawn in the upper left hand corner of the work area to identify the location of the current topic of conversation (e.g. a swing could represent the playground, a desk could represent the classroom).
- For the purpose of this research, the situations/topics to be discussed through a CSC will be those which relate to the target behaviours of interest.

Questioning:
Exploring the incident

Information is gathered about the situation to help 'complete the picture'.

1. Where are you? (child draw location symbol and a person)
2. Who else is here? (draws person/people)
3. What are you doing? (draws relevant items/actions)
4. What happened? What did others do? (draws relevant items/actions)
5. What did you say? (use talk symbol)
6. What did others say? (use talk symbol)
7. What did you think when you said that? (use thought symbol)
8. What did others think when you/they said/did that?

Sharing new insights

▸ The most difficult questions are often those concerning the thoughts, feelings and motivations of others.

▸ At natural 'opportunity' points, the perspective of the adult is shared with the child (e.g. if the child is struggling to answer a question, or a clear misunderstanding/misinterpretation is observed).

▸ The aim is to achieve a balance between gathering insights into the child's perspective, while sharing accurate social information.

▸ Never discredit the child's perspective even if they are incorrect in identifying what others think – instead, simply introduce another idea along the lines of e.g. "maybe your teacher/friend was thinking…. I'll write that in, maybe that's what they were thinking".

213

Sequence & Structure

- Comic strip boxes may be used to frame and sequence events (draw the boxes prior to the start of the conversation).
- If the student draws the event out of sequence, the situation is reviewed and numbers placed in each box to identify the order in which they occurred.
- If deemed useful, the boxes could be cut apart and placed in the correct order.
- Don't worry if the resulting CSC looks likes it would be confusing to people who had not been part of it – the key thing is that it makes sense to the student!

Concluding the Conversation

Summarise the conversation by reviewing the key points:
- If possible, the student may be encouraged to do this themselves. If not, try and encourage them to point to the drawings which illustrate events as they are summarised by the adult.
- The purpose of the summary is to 'pull together' the different aspects of the conversation prior to identifying new responses to the situation.

Lastly, new solutions to the situation are identified:
- Encourage the student to identify possible solutions/alternatives as far as is possible – the adult can supply ideas if necessary.
- Write down the ideas generated (there may be several).
- The pros and cons of each solution may be discussed (and could even be illustrated with drawings if that would help). Solutions which are no longer considered feasible are eliminated.
- You end up with a list of solutions which form a 'plan' for how the student could address the situation the next time it occurs.
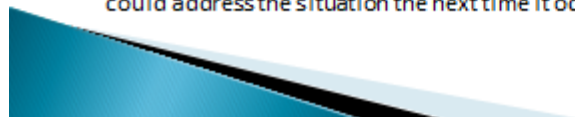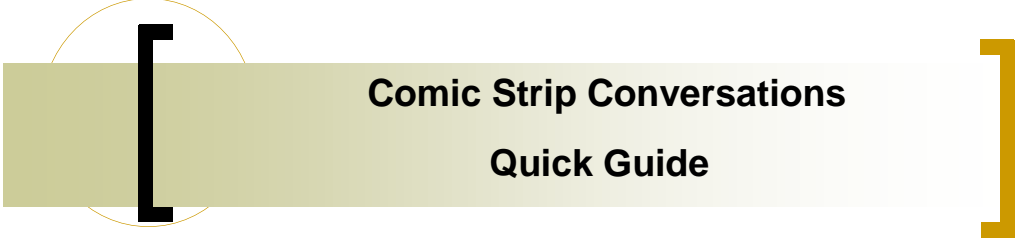
The use of colour

- Colour may be used to identify the emotional content of a statement, thought or question.
- Words could be written in different colours, or coloured pens could be used to underline words after they have been written.
- The **Colour Chart** lists suggested colours and associated feelings/motivations (but the student may wish to create their own).
- Colours can be introduced gradually over the course of several conversations – start with the basic ones.
- The student may make mistakes in identifying the appropriate emotions – if this happens, do not criticise their answer, but rather offer a second interpretation.
- More than one colour can be used to identify mixed feelings.

References

Gray, C. (1994). Comic Strip Conversations.

Comic Strip Conversations

Quick Guide

General points:

· **Encourage pupil to take the lead as much as possible with you 'guiding' the conversation (they talk, draw and write as much as possible)**

· **Place emphasis on considering what people may be *thinking***

· **Use strategies to assist in keeping a sequence and structure (e.g. boxes)**

1. Begin conversation with '**small talk**' before **introducing the topic** of conversation.

2. Draw **location symbol** (upper left-hand corner)

3. Gather information about the scenario through **questioning,** e.g.:

   *Who is here? (*draw people)
   What are you doing? (draw relevant items/actions)
   What happened? What did others do? (draw relevant items/actions)
   What did you say? (use talk symbol)
   What did others say? (use talk symbol)
   *What did you think when you said/did that? (*use thought symbol)
   What did others think when you/they said/did that? (thought symbol)

4. **Share your perspective** as and when appropriate

5. **Summarise** the conversation so far

6. Identify **new 'solutions'/alternative behaviours -** write or draw these (you may wish to discuss pros/cons of each idea).

7. **Action-plan formation** - decide which solutions/alternative(s) the pupil would like to try.

*In addition to the above process, it may be appropriate to introduce the pupil to **different symbols** (see symbol dictionary) and the **use of colour** to identify the emotional content of aspects of the situation.*

Intervention diary

Participant No:

Please complete each day that intervention activity is carried out (i.e. creation of a CSC or a review of a CSC).

Please note down any pertinent information regarding factors such as:

- the extent to which the pupil engaged with the activity and produced their own ideas
- the extent of adult guidance needed to think about the thoughts/feelings and in developing the action plan
- any amendments to the action plan following a review of the CSC
- anything else you feel is important when considering the pupil's response to the intervention.

Date	Action	Comments
	(i.e. 'CSC completed' or 'review of CSC')	

Appendix 14: Intervention Fidelity Checklist

Please complete this checklist following each Comic Strip Conversation that is completed.

Aspect of intervention	Present? (please put ✓ or x)	Comments
Adult and pupil sat next to each other		
Location symbol used to identify the setting of the topic of conversation		
Questioning used to identify *who* was involved in the situation		
Questioning used to identify *what was done and/or said* by the participant and relevant others		
Questioning used to identify the *thoughts* of the participant and relevant others		
The perspective of the adult was shared with the pupil as required		
Conversation was summarised		
Solutions/alternative behaviours to try in the future have been identified and recorded		
Key symbols - speech and thought bubbles - have been used		
The CSC has a clear structure/sequence		

The pupil was encouraged to play as active a role in the process as possible		
Additional features that may be present:		
Colour was used to visually illustrate emotional content		
Symbols other than speech/thought bubbles were used		

Appendix 15: Ethics committee approval letter

The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

SJ/wb
Ref: 635

Tuesday, 31 March 2015

Dear Joanne Page & Sarah Atkinson,

Ethics Committee Review

Thank you for submitting an account of your proposed research 'An investigation into the effectiveness of Comic Strip Conversations for improving the target social behaviours of primary-ages pupils on the autistic spectrum'.
That proposal has now been reviewed and we are pleased to tell you it has met with the Committee's approval.

However:

Please note the following comments from our reviewers;

The debrief for the child participants could be more age appropriate e.g. '….change some of your behaviours in school in a positive way' (will they know what this means)? The participant information sheet is much better in this respect. And is it really necessary to emphasise that they were observed unawares? 'This was done by watching ….' Could the debrief not simply read 'This was done by seeing how you got along at school in between times, to see whether activities like the Comic Strip Conversations might help you.'

Final responsibility for ethical conduct of your research rests with you or your supervisor. The Codes of Practice setting out these responsibilities have been published by the British Psychological Society and the University Research Ethics Committee. If you have any concerns whatever during the conduct of your research then you should consult those Codes of Practice. The Committee should be informed immediately should any participant complaints or adverse events arise during the study.

Independently of the Ethics Committee procedures, supervisors also have responsibilities for the risk assessment of projects as detailed in the safety pages of the University web site. Ethics Committee approval does not alter, replace, or remove those responsibilities, nor does it certify that they have been met.

Yours sincerely

Professor Stephen Jackson
Chair, Ethics Committee

Appendix 16: Tau-U outputs

Jack:

☐ J BL1	☐ J Int1	☐ J BL2	☐ J Int2	☐	☐	☐	☐	☐	☐
8.3 5 1.7 16.6 8.3	1.7 10 6.7 6.7 3.3 13.3 1.7 5 3.3	25 15 15 20 25 40 15 10 10	10 15 20 10 10						

Results

combine	to weighted	remove	clear all	download all

id	Label	S	PAIRS	TAU	TAUb	VARs	SD	SDtau	Z	P Value	CI 85%	CI 90%
trend:												
☐ 0	J BL1 vs J BL1	1	10	0.1000	0.1053	16.6667	4.0825	0.4082	0.2449	0.8065	-0.488<>0.688	-0.572<>0.772
☐ 2	J BL2 vs J BL2	-11	36	-0.3056	-0.3284	92	9.5917	0.2664	-1.1468	0.2515	-0.689<>0.078	-0.744<>0.133
phase:												
☐ -	-	-	-	-	-	-	-	-	-	-	-	-
corrected baseline:												
☐ 1	J BL1 vs J Int1	-13	45	-0.2889	-0.2989	225	15	0.3333	-0.8667	0.3861	-0.769<>0.191	-0.837<>0.259
☐ 3	J BL2 vs J Int2	-10	45	-0.2222	-0.2500	225	15	0.3333	-0.6667	0.5050	-0.702<>0.258	-0.771<>0.326

Daniel:

☐ D BL1	☐ D Int1	☐ D BL2	☐ D Int2	☐	☐	☐	☐	☐	☐
1 13 4 7 8 4 2	0 3 2 1 0 0 0 1 0 0	3 0 1 0 0 4 0 2 0 2 1 0	1 3 0 0 0 2						

Results

combine	to weighted	remove	clear all	download all

id	Label	S	PAIRS	TAU	TAUb	VARs	SD	SDtau	Z	P Value	CI 85%	CI 90%
trend:												
☐ 2	D BL1 vs D BL1	-2	21	-0.0952	-0.0976	44.3333	6.6583	0.3171	-0.3004	0.7639	-0.552<>0.361	-0.617<>0.426
☐ 3	D BL2 vs D BL2	-5	66	-0.0758	-0.0870	212.6667	14.583	0.2210	-0.3429	0.7317	-0.394<>0.242	-0.439<>0.288
phase:												
☐ 0	D BL1 vs D Int1	-68	77	-0.8831	-0.9007	487.6667	22.0832	0.2868	-3.0793	0.0021	-1.296<>-0.470	-1.355<>-0.411
☐ 1	D BL2 vs D Int2	-1	72	-0.0139	-0.0165	456	21.3542	0.2966	-0.0468	0.9626	-0.441<>0.413	-0.502<>0.474

Robert:

☐ R BL	☐ R Int	☐	☐	☐	☐	☐	☐	☐	☐
6 3 3 3 3 1 4 1 2	2 2 3 2 1 1 2 2 3 1								

Results

combine	to weighted	remove	clear all	download all

id	Label	S	PAIRS	TAU	TAUb	VARs	SD	SDtau	Z	P Value	CI 85%	CI 90%
trend:												
☐ 0	R BL vs R BL	-15	36	-0.4167	-0.4615	92	9.5917	0.2664	-1.5639	0.1179	-0.800<>-0.033	-0.855<>0.022
phase:												
☐ -	-	-	-	-	-	-	-	-	-	-	-	-
corrected baseline:												
☐ 1	R BL vs R Int	-24	90	-0.2667	-0.2981	600	24.4949	0.2722	-0.9798	0.3272	-0.659<>0.125	-0.714<>0.181

Owen:

☐ O BL1	☐ O Int1	☐ O BL2	☐ O Int2	☐	☐	☐	☐	☐	☐
7 9 12 4 7 4 3 3	2 2 2 0 1 0 0 2 2 1 0	15 3 13 9 5	6 2 6 3 4						

Results

combine	to weighted	remove	clear all	download all

id	Label	S	PAIRS	TAU	TAUb	VARs	SD	SDtau	Z	P Value	CI 85%	CI 90%
trend:												
☐ 0	O BL1 vs O BL1	-17	28	-0.6071	-0.6415	65.3333	8.0829	0.2887	-2.1032	0.0354	-1.023<>-0.191	-1.082<>-0.132
☐ 2	O BL2 vs O BL2	-4	10	-0.4000	-0.4000	16.6667	4.0825	0.4082	-0.9798	0.3272	-0.988<>0.188	-1.072<>0.272
phase:												
☐ -	-	-	-	-	-	-	-	-	-	-	-	-
corrected baseline:												
☐ 1	O BL1 vs O Int1	-71	88	-0.8068	-0.8068	586.6667	24.2212	0.2752	-2.9313	0.0034	-1.203<>-0.410	-1.260<>-0.354
☐ 3	O BL2 vs O Int2	-10	25	-0.4000	-0.4082	91.6667	9.5743	0.3830	-1.0445	0.2963	-0.951<>0.151	-1.030<>0.230

Gareth:

☐ G BL1	☐ G Int1	☐ G BL2	☐ G Int2	☐	☐	☐	☐	☐	☐
8 2 5 3 4 5 4 3	3 0 0 2 8 2 0 1	6 2 2 8 3 2 5 4 8 3	2 3 2 1 2 2						

Results

combine	to weighted	remove	clear all	download all

id	Label	S	PAIRS	TAU	TAUb	VARs	SD	SDtau	Z	P Value	CI 85%	CI 90%
trend:												
☐ 0	G BL1 vs G BL1	-5	28	-0.1786	-0.1887	65.3333	8.0829	0.2887	-0.6186	0.5362	-0.594<>0.237	-0.653<>0.296
☐ 2	G BL2 vs G BL2	6	45	0.1333	0.1412	125	11.1803	0.2485	0.5367	0.5915	-0.224<>0.491	-0.275<>0.542
phase:												
☐ -	-	-	-	-	-	-	-	-	-	-	-	-
corrected baseline:												
☐ 1	G BL1 vs G Int1	-38	64	-0.5938	-0.6179	362.6667	19.0438	0.2976	-1.9954	0.0460	-1.022<>-0.165	-1.083<>-0.104
☐ 3	**G BL2 vs G Int2**	**-46**	**60**	**-0.7667**	**-0.8679**	**340**	**18.4391**	**0.3073**	**-2.4947**	**0.0126**	**-1.209<>-0.324**	**-1.272<>-0.261**

Appendix 17: Visual analysis inter-rater reliability script

<u>Inter-Rater Questionnaire</u>

The table below describes the features that were considered in the visual analysis of the graphs:

Feature	Definition of the Feature
Level	The overall average (mean) of the measures within a phase. *In the present research, a <u>decrease</u> in mean level is indicative of a desirable intervention effect.*
Trend	The slope of the best-fitting straight line for the measures within a phase
Variability	The range, variance or standard deviation of the measures about the best-fitting line. *Low levels of variability are indicative of a more stable and reliable data set.*
Immediacy of effect	The change in level between the last three data points in one phase and the first three data points of the next *The more rapid the effect, the more convincing the inference that change can be attributed to the introduction of the independent variable.*
Overlap	The proportion of data from one phase that overlaps with data from the previous phase. *The smaller the proportion of overlap, the more convincing the demonstration of an effect.*

Please look at each of the graphs provided for each participant along with the corresponding descriptive summary table of the visual analysis. Then, consider the following question and record your responses with a tick on the recording sheet: **How certain or convinced are you that there was a practical, significant improvement in the participant's behaviour between the baseline and intervention phase?**

Please note that improvement would be demonstrated by a <u>decrease</u> in behaviour frequency.

<u>Inter-Rater Recording Sheet</u>

How certain or convinced are you that there was a practical, significant improvement in the participant's performance between the baseline and intervention phase?

Jack	1: Not at all certain	2: Uncertain	3: It is possible	4: Reasonably certain	5: Very certain
TB1					
TB2					

Daniel	1: Not at all certain	2: Uncertain	3: It is possible	4: Reasonably certain	5: Very certain
TB1					
TB2					

Robert	1: Not at all certain	2: Uncertain	3: It is possible	4: Reasonably certain	5: Very certain
TB1					

Owen	1: Not at all certain	2: Uncertain	3: It is possible	4: Reasonably certain	5: Very certain
TB1					
TB2					

Gareth	1: Not at all certain	2: Uncertain	3: It is possible	4: Reasonably certain	5: Very certain
TB1					
TB2					

Appendix 18: Cohen's Kappa output for inter-rater agreement of visual analysis

```
CROSSTABS
  /TABLES=Rater1 BY Rater2
  /FORMAT=AVALUE TABLES
  /STATISTICS=KAPPA
  /CELLS=COUNT
  /COUNT ROUND CELL.
```

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Asst of rater1 * Asst of rater 2	9	100.0%	0	0.0%	9	100.0%

Asst of rater1 * Asst of rater 2 Crosstabulation

Count

		Asst of rater 2			Total
		Not at all certain	Uncertain	It is possible	
Asst of rater1	Not at all certain	3	0	0	3
	Uncertain	1	1	0	2
	It is possible	0	1	2	3
	Reasonably certain	0	0	1	1
Total		4	2	3	9

Symmetric Measures

	Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement Kappa	.518	.207	2.447	.014
N of Valid Cases	9			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.