The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

The Impact of Cooperative Learning on Pupils' Peer Acceptance and Perceptions of Peer Support

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

Helen Smith

May 2017

Table of Contents

List of Tables

List of Figures

List of Acronyms

Acronym	Description
BPS	British Psychological Society
CL	Cooperative Learning
CSS	Classroom Supportiveness Scale
DV	Dependent Variable
EP	Educational Psychologist
EPS	Educational Psychology Service
NQT	Newly Qualified Teacher
PDG	Play Different Gender
PSG	Play Same Gender
SEGPA pupils	Pupils in French middle schools who have been placed in special classes based on their writing and reading levels (Section d'Enseignement Générale et Professionnelle Adaptee)
SENCO	Special Educational Needs Coordinator
SIS	Social Inclusion Survey
STAD	Student Teams-Achievement Divisions
TEP	Trainee Educational Psychologist
WDG	Work Different Gender
WoE	Weight of Evidence
WSG	Word Same Gender

Abstract

Cooperative Learning, an approach that emphasises positive interdependence and individual accountability in the classroom, has been widely promoted as a pedagogical approach that can enhance not only pupils' academic outcomes, but also their social outcomes (Johnson & Johnson, 2000). More specifically, research has indicated some potential for Cooperative Learning to impact on pupils' peer acceptance and perceptions of peer support. Questions remain, however, around a potential differential impact, and the extent to which findings can be generalised to UK, mainstream classes.

A quasi-experimental group design was used to investigate the impact of Cooperative Learning on pupils' peer acceptance and perceptions of peer support. Data were collected from four schools (intervention *N*=101; control *N*=102). Cooperative Learning was delivered by teachers in Year 4 or 5 classes over an 11 week period. A week of social skills training was included in the delivery.

The Social Inclusion Survey (Frederickson and Furnham, 1998) was used to measure peer acceptance across four contexts *(work same gender, work different gender, play same gender, play different gender),* and perceptions of peer support were measured using the Classroom Supportiveness Scale taken from the Sense of School Community Scale (Battistich, Solomon, Kim et al., 1995).

Findings indicate that Cooperative Learning can enhance pupils' perceptions of peer support. There appeared to be a differential impact of the intervention, in favour of pupils with initially low levels of perceived peer support.

Findings also indicate that Cooperative Learning can enhance pupils' peer acceptance, particularly in the work context. Findings for the play context were mixed, with statistically significant increases found in different gender but not same gender peer acceptance ratings. There appeared to be a differential

impact of the intervention, in favour of pupils with initially low or medium levels of peer acceptance.

Although some caution is advised in interpreting results – particularly since the degree of impact was not consistent across individual classes – it is concluded that Cooperative Learning approaches can have a valuable impact on pupils' peer acceptance and perceptions of peer support, and that, reassuringly, they appear to have the greatest impact on those pupils with initially lower levels of peer acceptance and perceived peer support.

Acknowledgements

For all their support throughout my doctoral training course, I would like to thank my wonderful friends and family, particularly my parents, my partner and my children. I could not have done it without so much support and encouragement. Thank you all from the bottom of my heart.

I would also like to thank the tutor team at the University of Nottingham, with particular mention to Nathan Lambert for his unwavering support and belief in me, and to Anthea Gulliford for carefully supporting my personal and professional development. It truly has been a journey, and I am a better person for it. Thank you to all the staff who supported me along the way.

Thanks also to my placement EPS, where I was granted flexibility and autonomy to develop my practice and complete my research, and to the school staff who worked so hard in delivering Cooperative Learning.

And last, but definitely not least, thanks to Team TEP. I have been incredibly lucky to be part of such a close and supportive cohort. Thank you for sharing in the ups and downs, and the many laughs along the way. You have made the experience one which I will always cherish.

Introduction

There exists a wealth of literature pointing to a range of positive psychological outcomes associated with both peer acceptance and peer support - as well as to the complexities of intervening effectively to support children who are having difficulties in these areas (Bierman, 2004).

These were issues encountered by the author during her experience as a teacher and as an assistant psychologist, and again during professional placement as a TEP, with schools frequently asking for strategies to support children, or groups of children, who were having difficulties in gaining peer acceptance, and in eliciting and maintaining peer support.

During professional training, the author became aware of 'Cooperative Learning' (CL) as a pedagogical approach (Johnson & Johnson, 2000). CL has been defined by Johnson & Johnson (2008) as 'the instructional use of small groups so that students work together to maximize their own and each other's learning' (p.26), with the emphasis on positive interdependence and individual accountability distinguishing CL from generic group work.

With suggestions in the literature that CL could potentially facilitate increases in peer acceptance and perceptions of peer support on a universal, whole class level without the need for extensive training, expensive materials, significant time commitment or withdrawal groups, the author believed that it was worthy of further investigation. However, the CL evidence base, whilst indicating potential, was limited, particularly with regards to its impact on perceptions of peer support. With regards to enhancing peer acceptance, the evidence base lay primarily outside of the UK, and often focused on special populations, or involved a high level of training for teachers.

Thus, the principal aim of this research was to investigate the impact of CL on peer acceptance and perceived peer support of children in UK mainstream primary classes. In so doing, the study aimed to build on the CL and peer acceptance literature. It also sought to make a unique contribution with regards to the impact of CL on pupils' own perceptions of peer support.

Furthermore, a small number of studies have suggested the possibility that CL may not necessarily bring universal benefit in enhancing children's peer acceptance. Indeed, there are some suggestions that the least accepted pupils may remain particularly stable in their status (Jules, 1991), or become *more* rejected by their peers following CL (Garel, 2001).

Thus, a second focus of the research was to explore the impact of CL on children with differing initial levels of peer acceptance and perceptions of peer support, before focusing more specifically on those children with initially low levels. In so doing, the study sought to further understand the impact of CL on the population with whom an EP may typically work, and to provide a further unique contribution to the peer acceptance and peer support literature.

1. Literature Review

This chapter outlines the key theoretical and empirical basis for the research, discussing literature surrounding peer acceptance, perceptions of social support, and Cooperative Learning. This is followed by a systematic review of the literature regarding the impact of CL on peer acceptance and perceptions of peer support. The research questions are presented, and the rationale for these described.

1.1 Social Inclusion and Social Exclusion: A Broad Perspective

Social inclusion:	‘‘the act of making all groups of people within a society feel valued and important’
Social exclusion	‘‘the act of making certain groups of people within a society feel isolated and unimportant’

Source: Collins English Dictionary online (Collins, 2017)

Abrams, Hogg, & Marques (2005) argue that social inclusion, or indeed exclusion, can be viewed at a transnational and societal level, all the way down to inter and intra group and personal levels. With a universal desire to interact, be connected to, and experience caring for others (Baumeister & Leary, 1995), exclusion at all levels can lead to conflict, inequality and mistrust (Abrams, Hogg, & Marques, 2005).

Sen (2000) notes that a key feature of exclusion is the inability to interact freely with others, which he describes as a deprivation in itself, as it can both directly and indirectly impoverish a person’s life, creating further deprivation, further exclusion, and further limiting opportunity.

As the meaning of social inclusion or exclusion can be viewed as contextually contingent, it is important that the construct is understood within the boundaries of a specific context (Silver, 2007).

For the purposes of this research, the context is that of the school, a setting which contains a framework of relationships in which pupils seek inclusion, acceptance, support and belonging (Abrams, Hogg, & Marques, 2005). The focus on social inclusion within the school context will be through two primary indicators as highlighted by Symes & Humphrey (2010):

- peer acceptance
- perceptions of peer support

1.2 Peer Acceptance: Social Inclusion in the School Context

One indicator of social inclusion within a school context is that of peer acceptance. This section will define peer acceptance, before considering the importance of acceptance by the peer group in social, emotional, academic, and long term outcomes.

1.2.1 A Definition of Peer Acceptance

Peer acceptance and rejection can be defined as 'the degree to which members of a particular peer group…like or dislike a child' (Bierman, 2004, p.7).

Peer acceptance is generally viewed as conceptually distinct to the notion of friendship (Gest, Graham-Bermann, & Hartup, 2001), as while the skills which underlie peer acceptance are likely to also be important in the development of supportive and strong friendships, the former reflects the perspective of the

child's peer group, whereas the latter reflects voluntary, dyadic relationships (Gifford-Smith & Brownell, 2003).

1.2.2 The Importance of Peer Acceptance

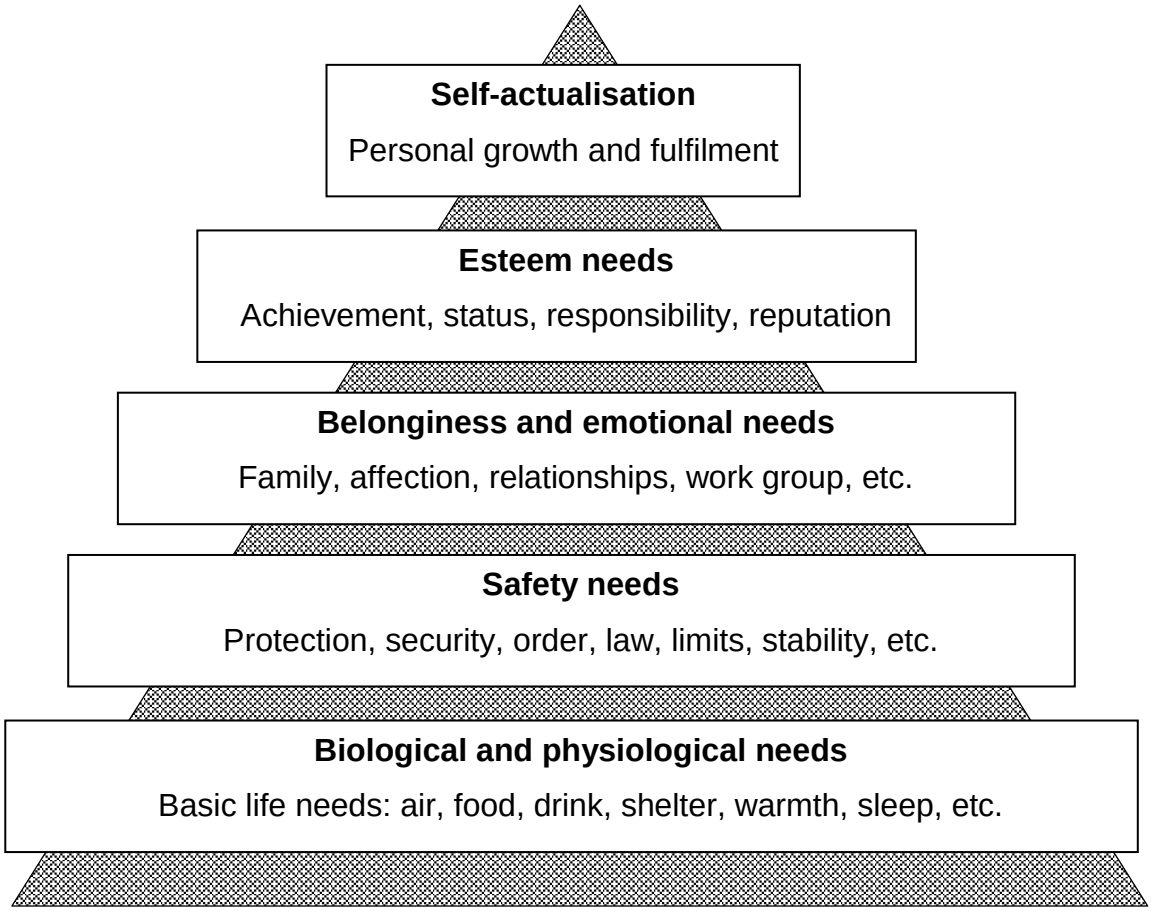
By the time they are in school, children mix with large peer groups and social relationships are an important part of their life (Slater & Bremner, 2011). With primary classrooms relatively self-contained, group level factors including being liked and accepted or disliked and rejected are central in shaping children's social experiences, and their social inclusion or exclusion (Rubin, Begle, & McDonald, 2012). The importance of peers and the peer group continues to increase throughout childhood and into adolescence (Coleman & Hendry, 1999). Yet literature indicates that at least 10% to 16% of school-aged children are rejected by their peers (Bierman, 2004; Terry & Coie, 1991).

1.2.2.1 The Need to Belong

In their 'belongingness hypothesis', Baumeister & Leary (1995) argue that the need to belong is one of the most important human needs in life, with a universal desire for individuals to have contact with people to whom they feel connected, leading people to seek, develop and maintain frequent and fulfilling interactions with others (Levett-Jones, Lathlean, Maguire, & McMillan, 2007). While peer acceptance can grant a sense of belonging (Brown, Mory, & Kinney, 1994), a lack of acceptance and a lack of belonging is often viewed as maladaptive, with rejection and exclusion leading to powerful negative feelings (Baumeister & Leary, 1995).

In Maslow's (1970) hierarchy of needs (Figure 1.1), feelings of belonging are placed before those of esteem and self-actualisation. As Maslow theorised that progression through the hierarchy could only be achieved once the lower stage needs were met, a sense of belonging is viewed as important as a stepping stone towards further wellbeing.

Figure 1.1 Maslow's hierarchy of needs



adapted from Jolliffe (2007, p.2)

With a fundamental need to belong and be accepted, it is perhaps unsurprising that literature points to the importance of experiences with peers and the peer group in all facets of a child's life, including the development of their cognitive, academic, social, and emotional skills (Rubin, Bukowski and Parker, 2006). These aspects will now be discussed further.

1.2.2.2 Peer Acceptance and Social and Emotional Development

Peer relationships form an important developmental context for children (Cillessen & Mayeux, 2004), providing them with an opportunity to practice and master critical social skills (Bierman, 2004).

Erwin (1998) writes that it is with their peers that children are most likely to experience conflicting ideas and explanations. This provides children with a natural platform to develop their negotiation and conflict management skills, as well as the skills required to work cooperatively, to manage communication, and to build and maintain close relationships. Children who are accepted by their peers are more likely to have ready access to positive peer interactions and positive models that support them in developing a range of prosocial behaviours, shaping their capacities to interact comfortably with others, to control aggression, and to experience intimacy (Hartup, 1989).

On the other hand, children who are less well accepted by their peer group are likely to have a reduced and impoverished context in which to develop these critical social skills. Literature suggests that children with low peer acceptance are at an increased risk of negative developmental outcomes (Bierman, 2004), with peer rejection associated with both externalising and internalising behavioural difficulties (Herbert, 2002). Baumeister & Leary (1995) suggest that social, emotional and behavioural difficulties may arise from children's increasingly desperate attempts to establish and maintain relationships and gain a sense of belonging, and from strong feelings of frustration and powerlessness when they are unable to do so.

Literature also points to the importance of peer interactions and peer acceptance in promoting a sense of identity and self-worth (Barrett & Randall, 2001)**,** with children with higher levels of peer acceptance tending to report more positive perceptions of self-worth than children with lower levels of peer acceptance (Harter, 1996)**.**

1.2.2.3 Peer Acceptance and Educational Experiences

As well as guiding social and emotional development, higher levels of peer acceptance have been associated with increased school success, including greater levels of academic performance (Diehl, Lemerise, Caverly, Ramsay, &

Roberts, 1998), school adjustment (Buhs, Ladd, & Herald, 2006) and academic motivation (Furrer & Skinner, 2003).

Conversely, research indicates that peer rejection, both past and present, is associated with lower school attendance (DeRosier, Kupersmidt, & Patterson, 1994), lower academic achievement (Wentzel & Caldwell, 1997) and an increased risk of school dropout (Coie, Dodge, & Kupersmidt, 1990).

A motivational perspective suggests that positive peer relationships and feelings of peer acceptance may engender a sense of belonging and cohesion, which then motivates academic involvement, engagement, and interest (Juvonen, 2007). Similarly, peer acceptance might allow pupils more ready access to resources which promote achievement, for example, help from a peer and sharing of ideas (Wentzel & Caldwell,1997). Indeed, pupils who experience support from their peers, as well as from their teachers, are more likely to attach greater value to learning in school (Roeser, Eccles, & Sameroff, 1998). The importance of peer support, and perceptions of support, will be discussed further in section 1.3**.**

On the other hand, negative peer experiences may undermine motivation to attend school (Buhs & Ladd, 2001), perpetuating and exacerbating any academic difficulties. Buhs, Ladd, & Herald (2006) found that peer rejection was associated with decreased classroom participation, which appeared to be an antecedent to lower academic achievement. Furthermore, it is possible that experiences of peer rejection may increase anxiety, and interfere with concentration as well as with acquisition and retention of information (Nansel et al., 2001).

Whilst many commentators highlight a degree of causal influence of peer acceptance on social, emotional, and academic outcomes, whereby early peer problems may lead to ongoing interpersonal difficulties, school avoidance and withdrawal, it is likely that there is a bidirectional influence, with behavioural and individual characteristics also playing a part in sociometric status (Wentzel, 2003). Furthermore, the role of non-causal, confounding factors which may be

associated with peer problems and academic outcomes must also be acknowledged, for example the potential impact of early parenting experiences, attentional difficulties, and socioeconomic status (Woodward & Fergusson, 2000).

1.2.2.4 Peer Acceptance and Longer Term Outcomes

Peer acceptance appears to be moderately stable over the long term (Herbert, 2002), and is described by Hay (2005) as a good predictor of later peer acceptance and peer relationships. A study by Zettergren (2005) found that children initially identified as rejected at ten years old maintained their low peer acceptance status throughout the five-year study.

Longitudinal studies indicate that children who have difficulties in gaining peer acceptance may be more likely to experience difficulties in later life (Bierman, 2004). Peer rejection in childhood has been particularly associated with a range of externalising problems in adolescence, including delinquency, conduct disorder, attentional difficulties, and substance abuse (Kupersmidt & Coie, 1990), even when prior levels of aggression and externalising problems are controlled for (Ladd & Burgess, 2001). Woodward & Fergusson (2000) found that, on average, children with low peer acceptance received fewer academic qualifications and were more likely to be unemployed upon leaving school.

Although the mechanisms and models underlying these links are unclear, there is a suggestion that children's chronic peer relationship experiences, rather than simply their own individual social behaviours, are significant antecedents of later maladjustment (Ladd & Troop-Gordon, 2003). That is, the relational process of peer rejection itself may have its own debilitating consequences on children's psychological adjustment and their developmental course (Ladd, 2006). It may be that the stress of peer rejection leads to an increase in antisocial behaviour, or it may be that exclusion from the peer group reduces a child's access to this important context for social learning, as well as leading them to develop negative expectations about future interactions (Dodge et al., 2003).

1.3 Perceived Peer Support: Social Inclusion in the School Context

A further indicator of social inclusion within a school context is that of perceived peer support (Symes & Humphrey, 2010). This section will define social support, before focusing more specifically on perceptions of peer support.

1.3.1 A Definition of Social Support

Social support is a multidimensional and complex phenomenon, and one which lacks a definitive definition (Williams, Barclay, & Schmied, 2004). However, one popular definition comes from Saylor & Leach (2009, p.71) who write that social support is 'the *perception* [emphasis added] one has of being cared for, valued, and included by others'.

This support can be gained from multiple sources, including parents, teachers, classmates, close friends, and the wider community (Demaray, Malecki, Davidson et al., 2005).

In line with the definition offered by Saylor & Leach (2009), literature indicates that it is a person's subjective *perceptions* of available support, that is, the notion that others will be available if needed, and the individual's felt sense of being supported, rather than actual support, which is most closely linked to positive psychological outcomes (Wethington & Kessler, 1986).

It is these outcomes which will now be discussed.

1.3.2 The Importance of Perceived Social Support

Literature regarding perceived social support highlights associations with a wide range of psychological outcomes (Pisula & Lukowska, 2011).

In comparison to those with lower levels of perceived support, children and adolescents with higher levels of perceived support have been found to have more positive alternatives for solving problems and conflicts (Malecki & Demaray, 2003), higher self-concept scores, higher adaptive skills, and less externalising behaviour problems (Demaray & Malecki, 2002)**,** as well as more pro social behaviours (Malecki & Elliott, 1999)**.**

Studies have found lower levels of perceived peer support to be related to a range of emotional and behavioural problems, including higher levels of anxiety, social stress, and depression (Demaray & Malecki, 2002). Other studies note associations with poorer self-concept (Wenz-Gross, Siperstein, & Untch, 1997), lower pupil reports of pro social behaviours (Demaray & Elliott, 2001), and poorer adjustment during the transition to high school (Hirsch & DuBois, 1992).

One framework for understanding the potential route by which perceptions of social support can affect wellbeing is through the 'stress buffering' model. This suggests that simply the perception that others are available to help, if and when needed, increases an individual's ability to cope with demands, and means that they are more likely to appraise potentially threatening situations as less stressful. When a person perceives social support, stress felt from any given situation is reduced, and so too are any associated negative impacts (Lakey & Cohen, 2000).

A second framework takes a cognitive perspective in suggesting that a person's internal working models help them to anticipate the supportiveness of others, as based on their beliefs about their own self-worth and beliefs about the general availability and responsiveness of others. At the same time, these beliefs affect how they view the world and the actions of others. Indeed, Lakey & Cohen

(2000) note that people with high levels of perceived support tend to interpret the same behaviours as more supportive, and pay more attention to supportive behaviours.

Whilst Johnson, Johnson, Buckman, & Richards (1985) write that social support is a critical aspect of classroom life, it is an aspect which does not necessarily appear readily accessible to all pupils. A study by Demaray & Elliott (2001) found that although all children perceived the importance of peer support equally, those with 'ADHD type behaviours' reported lower perceptions of peer support than their classmates. Similarly, pupils with learning difficulties perceived less peer support and more stress than those without learning difficulties (Wenz-Gross & Siperstein, 1998).

Although the bulk of literature regarding perceptions of peer support has focused on adolescent populations, and, as with peer acceptance, directionality is likely complex, there does appear to be a case that both peer acceptance and perceptions of peer support are associated with psychological wellbeing, a sense of belonging, and a range of pro social attitudes and behaviours (Johnson, Johnson, & Anderson, 1983).

Focus will now turn to understanding the ways in which children may have more, or less, success in gaining peer acceptance and peer support.

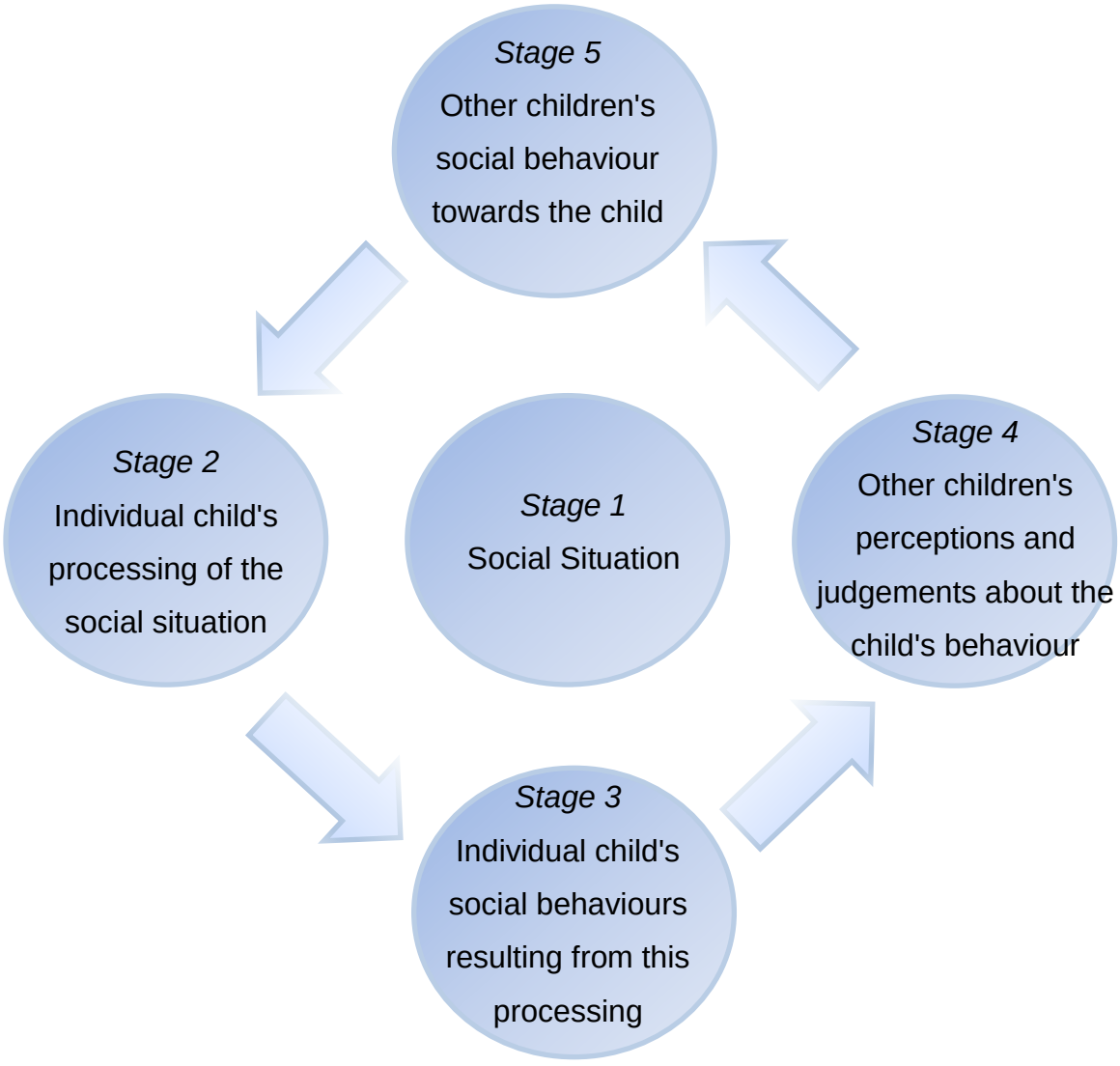
1.4 Factors Influencing Peer Acceptance and Peer Support: A Model of Social Exchange

Through an examination of the Dodge, Pettit, McClaskey, Brown, & Gottman (1986) model of social functioning and social exchange, the following section aims to consider the many factors influencing peer acceptance and peer support.

The Dodge et al. (1986) model describes a cyclical relation between social behaviour and social information processing, highlighting the complex

processes and reciprocal interactions which are likely to contribute to social acceptance or rejection, and access to peer support (Figure 1.2).

Figure 1.2 A model of social exchange in children (adapted from Dodge et al., 1986)



In considering a child's acceptance or rejection as the result of an interaction between individual and environmental factors within a certain social situation at a certain time (stage 1), the model helps in understanding the multiple processes which may establish, maintain and exacerbate the social and interpersonal difficulties of children who are not well accepted by their peer group (Price & Dodge, 1989).

A closer look at the role of an individual child (stages 2 and 3) and the child's peer group (stages 4 and 5) now follows.

1.4.1 Individual Characteristics

When processing the social situation (stage 2), the child's perception and understanding of the situation is central (Frederickson & Cline, 2009). The child needs to judge whether the peer approaching them is being friendly or aggressive, and therefore whether an appropriate response would be to run away, engage, or ignore (Smith, Cowie, & Blades, 2011).

This processing is affected by a child's past interpersonal experiences, which guide their perceptions, goals, and expectations in interactions (Bierman, 2004). Literature suggests that children who are rejected by their peers display inadequate and biased processing of social cues and overly negative attributions of peers' intentions (Price & Dodge, 1989), being more likely to attribute hostile intent to others (Crick & Dodge, 1994), and to generate aggressive solutions to social situations (Crick & Ladd, 1990).

On the other hand, studies suggest that well accepted children are more accurate in their encoding of social cues (Dodge & Price, 1994), and more likely to favour relational rather than instrumental goals (Nelson & Crick, 1999). These children are more likely to report higher perceptions of social support, perhaps due in part to their higher likelihood of interpreting ambiguous social situations as pro social and supportive, rather than aggressive (Anan & Barnett, 1999), and their tendency to pay more attention to supportive behaviours (Lakey & Cohen, 2000).

The behaviour selected at stage 3, in line with their earlier processing of the social situation, plays a critical role in forming and maintaining a child's social status (Gifford-Smith & Brownell, 2003).

Although there does not appear to be a single profile of a rejected child, Bierman (2004) argues that most poorly accepted children show at least one of the following 'social toxin' behaviours:

- low rates of pro social behaviour.
- high rates of aggressive/disruptive behaviour.
- high rates of inattentive/immature behaviour.
- high rates of socially anxious/avoidant behaviour.

Children who encounter repeated failure in meeting their social goals can develop negative thoughts and feelings about themselves (Nelson, Rubin, & Fox, 2005). Through anxiety, anger, or fear, these emotions can lead to aggression or withdrawal, and further rejection, and towards patterns of thinking and behaving that exacerbate their social difficulties (Toner & Munro, 1996). This can lead to a cycle of deprivation, whereby a child is unable access peer interactions that would support them in strengthening their social skills and confidence, and further excludes them from accessing peer support.

East, Hess, & Lerner (1987) found that adolescent girls who were rejected by their peers perceived lower levels of available peer support than their classmates. They suggested that the impaired social skills which may have led to peer rejection may also have inhibited the girls from receiving social support from peers, and that this appeared to be compounded by a tendency for the girls to increasingly withdraw from peer interactions, further diminishing their opportunities for receiving support.

1.4.2 The Role of Peers

The way in which a child's social behaviour is interpreted by their peers (stage 4) will have a crucial impact on the extent to which they are accepted or rejected by their peers (Frederickson & Cline, 2009). Literature suggests that once a child's social reputation is established, the social cognitive processing and social behaviour of the individual members of the peer group may act to maintain and perpetuate this (Hymel, Wagner, & Butler, 1990).

A child's reputational status biases the ways in which peers perceive and interpret that child's behaviour, and in how they subsequently respond to them (Hymel, Wagner, & Butler, 1990). Stereotypic processing by a peer group can mean that there is a tendency for children to selectively attend to inappropriate behaviour from rejected children, and overlook incidences of pro social behaviours (Hymel, Wagner, & Butler, 1990). The behaviour of less accepted children is more likely to be interpreted in a negative manner, compared to the behaviour of more accepted children who are viewed as more responsible for their positive and less responsible for their negative behaviours (Waas & Honer, 1990). Negative reputational biases can be insurmountable obstacles for many poorly accepted children, and may contribute to self-fulfilling prophecies (Bierman, 2004).

Moreover, peers themselves are more likely to behave aggressively towards rejected children (Bierman, 2004). Indeed, studies suggest that rejected children are subjected to a higher number of aversive interpersonal interactions in school than non-rejected children (Zakriski & Coie, 1996), and are more likely to be teased for non-normative or unusual behaviour (Gifford-Smith & Brownell, 2003). For children who have low peer acceptance, peer responses can evoke, maintain, and escalate their problematic social behaviours (Bierman, 2004).

Furthermore, peers can control the niches of social opportunity available to other children (Bierman, 2004). For poorly accepted children, peers can become increasingly unavailable and unresponsive, reducing opportunities for the child to access the high-quality interactions and rich social learning experiences which would model and support their social interactional skill development (Ladd & Asher, 1985). Compared to their well accepted peers, less well accepted children tend to interact with fewer peers, younger peers, and have more low quality exchanges which involve high levels of conflict (Bierman, 2004). Children who experience peer rejection are likely to be isolated from their peers, and therefore have limited avenues for seeking social support (Lanier, 2007). As Ladd (1983) notes, the children who would benefit

most from being members of groups with socially skilled peers are the least likely to be admitted.

1.5 Intervening with Peer Acceptance and Peer Support

As highlighted in the previous sections, the processes contributing to peer acceptance are complex, with many potential barriers for poorly accepted children in changing their status.

Indeed, studies have found that a child's peer group status is moderately stable over the long term (Herbert, 2002), with peer ratings of children's desirability as a work partner stable across a school year (Putnam, Markovchick, Johnson, & Johnson, 1996). Similarly, research indicates that social support is stable within a given developmental period, leading Demaray, Malecki, Davidson et al. (2005) to suggest that significant changes in levels of perceived peer support within an academic year are unlikely.

With peer rejection an interactional process, intervention approaches which recognise the role of the social context in maintaining patterns of social behaviour and social status (Erwin, 1998), and therefore focus on altering these social structures to ensure a receptive and facilitative peer environment - as well as attending to the rejected child - are crucial (Price & Dodge, 1989).

Frederickson & Cline (2009) write that effective interventions for children who are having difficulty gaining peer acceptance need to involve situations which allow:

- rejected children and their peers to interact in positive ways
- opportunities for prejudices to be reduced
- everyone to make a valued contribution, with differences in competence minimised
- rejected children to learn new skills

- opportunities for peers to change their interpretations of, and responses to, the rejected child's behaviour

In a similar vein, Sheldon (2004) suggested that social supportiveness interventions should seek to:

- improve individual social skills
- build stronger ties to peers
- reduce negative interactions

Finally, Johnson & Johnson (2013) note that if children are to perceive that they are liked, supported and accepted, then they must not only be taught appropriate skills, but must also have the opportunity to use them.

Cooperative Learning (CL) is an approach which appears to hold the potential to address all these aspects of effective intervention, and potentially impact on both peer acceptance and perceptions of peer support through addressing both a child's individual social skills as well as the peer group context (Johnson & Johnson, 1999).

1.6 Cooperative Learning

CL is not a new idea (Marr, 1997), and indeed, has been touted as one of the most widely researched instructional techniques (Johnson & Johnson, 1999).

Its modern day roots may be traced back to the opening of the 'Lancastrian' school in 1806, which emphasised CL in order to support the integration and socialisation of pupils from a range of cultures and backgrounds (Marr, 1997). In the late 1800s, Colonel Francis Parker advocated for CL in his Massachusett schools, believing that pupils could reach their full capacity if shared learning rather than competition was the main motive of school tasks (Marr, 1997). A later advocate in the 1920s was Dewey (1966), who emphasised the social aspects of learning, and the importance of schools in fostering interpersonal communication and group involvement (Gillies & Ashman, 2003).

Today, there are a range of approaches to CL, built on some core underlying principles. The principles and approaches, as well as theoretical underpinnings, will be described in the coming sections.

1.6.1 What is Cooperative Learning?

'Cooperative learning is the instructional use of small groups so that students work together to maximize their own and each other's learning' (Johnson & Johnson, 2008, p.26).

There are five key elements highlighted by Johnson, Johnson, & Holubec (1994) as core in effective CL, and which distinguish CL from traditional group work. These are:

Positive interdependence: the 'sink or swim together' feeling which occurs when group members perceive that they cannot succeed unless everyone succeeds. Positive interdependence has been described as the heart of CL (Thomson & Brown, 2000).

Individual accountability: where there can be no 'hitchhiking' or social loafing. Instead, each pupil is accountable for their own share of the work, and contribution to the group.

Face to face promotive interaction: occurring when pupils work in small groups in close proximity to each other (Andre, Louvet, & Deneuve, 2013). In this way, CL facilitates the development of a peer support system, within which group members share resources, support, encourage, and praise each other (Johnson & Johnson, 2013).

Interpersonal and small group skills: the explicit, systematic teaching of interpersonal and small group skills to all children is central in CL, as structuring a task to create positive interdependence is insufficient if pupils do not have the

appropriate skills and strategies for interacting with one another (Johnson, Johnson, Stanne, & Garibaldi, 1990).

CL Literature indicates that, across all ability levels, pupils who have received training in small group social skills are more inclusive of others, more cooperative, more involved in group activities, more respectful and considerate of the suggestions of others, more responsive to the needs of their peers, and provide more help and richer explanations to each other, when compared to pupils who have not received such initial training (e.g. Gillies, 2016).

Group processing: pupils reflect on and evaluate how well they worked together, considering what factors and actions were helpful and unhelpful, and how to build on this next time to further improve group functioning (Thomson & Brown, 2000).

Johnson, Johnson, Stanne, & Garibaldi (1990) reported higher achievement gains for pupils who had received group processing opportunities when compared to those who did not. Similarly, Bertucci, Johnson, Johnson, & Conte (2012) found that pupils who received CL with group processing gained more marks on an achievement test than pupils who received CL without group processing.

Having identified the key facets of CL, theories underpinning the approach will now be discussed.

1.6.2 Theories Supporting and Underpinning Cooperative Learning

Johnson & Johnson (2008) point to a close relationship between theory, research and practice as being one of the most distinctive characteristics of CL.

Three main theories cited in relation to CL are:

- social interdependence theory
- cognitive developmental theory
- behavioural-social learning theory

These will now be discussed in further detail.

1.6.2.1 A Social Interdependence Perspective

Social interdependence theory, derived from the work of Lewin (1935) and Deutsch (1949), is based on the understanding that goal specification determines how individuals act and interact, which in turn determines the resulting outcomes, such as achievement and interpersonal liking (Peterson, Johnson, & Johnson, 1991).

Deutsch (1949) identified two types of social interdependence. Positive interdependence and cooperation occurs when individuals perceive that achieving their goals is inextricably linked to the others in the group reaching theirs. This creates an enhanced motivation to engage in promotive interaction, that is, to help each other to reach shared goals (Johnson & Johnson, 2009). On the other hand, negative interdependence and competition occurs when only one or a few individuals can attain a goal, meaning that the motivation is for oppositional interaction whereby pupils discourage and obstruct others' efforts.

Deutsch (1949) also described situations in which there is an absence of goal interdependence, which results in individualistic effort. In these situations, where an individual's achievement of his or her goals is unrelated to the efforts and achievement of peers, then there will be limited interaction (Johnson & Johnson, 2009).

In CL, the structuring of positive interdependence is viewed as key in creating promotive interactions and positive outcomes, including the development of positive peer relationships (Johnson & Johnson, 2013).

1.6.2.2 A Cognitive Developmental Perspective

Developmental Perspective:

From a developmental perspective, the theories of Vygotsky (1978) and Piaget (1959) both have implications for the processing mechanisms underlying successful CL (Dansereau & Johnson, 1994).

The work of Vygotsky (1978) highlights knowledge as social, with learning taking place with, and alongside, others during cooperative efforts (Johnson, Johnson, & Holubec, 1994). By working cooperatively with a heterogeneous group of peers, pupils encounter new knowledge and ways of thinking (O'Connor & Jenkins, 2013), allowing them to progress through their zone of proximal development (Jolliffe, 2007).

Similarly, Piaget (1959) highlighted the importance of peer interaction in learning, pointing to the role of confrontation and conflict as catalysts for change and cognitive growth (Johnson & Johnson, 1999). Through interaction and working cooperatively with their peers, discussions lead to cognitive conflicts. As these are resolved, and inadequate reasoning is exposed, learning and progress occurs (Johnson, Johnson, & Holubec, 1994).

Cognitive Elaboration Perspective:

Related to the developmental perspective, is the cognitive restructuring or cognitive elaboration perspective, whereby an individual is required to cognitively rehearse and restructure information in order to incorporate it in to existing cognitive structures and retain it in their memory (Wittrock, 1989). CL

provides the context and opportunity for learners to consolidate knowledge by explaining, paraphrasing and revising information (Marr, 1997).

1.6.2.3 A Behavioural-Social Learning Perspective

Based on the work of Skinner (1953), a behavioural learning approach focuses on the impact of group reinforcers and rewards on learning, with the assumption that actions which are rewarded extrinsically will be repeated. Indeed, motivational perspectives and Slavin's (1980) work emphasises the importance of group rewards in motivating members of a CL group.

Bandura's (1971) work on social learning theory highlights observational learning and modelling as powerful forces for learning, as pupils look to emulate those who are successful within their CL group (Kagan & Kagan, 2009).

In line with behavioural-social perspectives, CL provides a social interaction situation where pupils can apply, practice and refine their skills in real world contexts, receiving immediate, supportive, and frequent corrective feedback from their peers, which makes the feedback highly desirable (Kagan & Kagan, 2009).

1.6.2.4 A Summary and Overview of the Perspectives

Johnson, Johnson, & Holubec (1994) note that there are fundamental differences in the basic assumptions held within each perspective.

For example, when understanding the gains reportedly made through cooperative working, social interdependence theory considers the importance of joint aspirations, group cohesiveness, and shared goals in rousing cooperation, whereas the behavioural-social learning perspective looks towards the role of extrinsic motivators and rewards in inspiring cooperation. Equally, whereas a cognitive developmental perspective focuses on individual factors occurring

within the individual, the social interdependence perspective takes a more relational approach.

However, rather than arguing for a ‘correct’ perspective, Slavin (2004) argues that the different theoretical perspectives should be viewed as complementary rather than contradictory, with each perspective contributing to a more complex understanding of how CL may affect pupil outcomes at different points, depending on the timing, context, and task.

1.6.3 Models of Cooperative Learning

Just as there are different theoretical perspectives on cooperation, so too there are different ways of operationalising and implementing CL in the classroom (Peterson, Johnson, & Johnson, 1991). Indeed, CL is essentially a generic term, covering a variety of methods, ranging on a continuum from concrete and prescribed approaches to more conceptual and flexible approaches (Johnson & Johnson, 2002).

Three major approaches to CL will now be described.

1.6.3.1 Johnson and Johnson’s ‘Learning Together’ Conceptual Approach

Described as one of the more popular models of CL, the Johnson and Johnson learning together approach is conceptually based and highly flexible (Murphy, Grey, & Honan, 2005). The approach provides suggestions of how to implement each of the five essential elements (as highlighted in section 1.6.1), with teachers able to select the ways in which they do so. Pupils work in small, heterogenous groups, on activities which are selected in order to maximise positive interdependence and individual accountability, for example, through goal interdependence which requires a single group product, or by pupils being asked to teach someone else what they have learned.

1.6.3.2 Kagan's Structural Approach

The Kagan structural approach uses a number of structures to support CL activities.

A structure is defined as a 'content free way of organizing social interaction in the classroom' (Kagan, 1990, p.12). The structures are designed to incorporate the key elements of CL. Each structure has several steps, and can be used by teachers to convert their existing practices and lesson content into CL activities in a range of different subject areas and across different age ranges.

Kagan (1990) argues that this approach is supportive to teachers, as to master CL as an abstract entity is a difficult task, but to learn and apply a small number of structures is a relatively easy undertaking. Structures may be used on their own, or in succession during a lesson, in order to best match intended learning outcomes and teaching objectives. With the knowledge of just five CL structures, Kagan (1990) writes that a teacher could meaningfully include CL in all lessons all year.

The Kagan approach highlights four basic principles using the acronym 'PIES' (Kagan & Kagan, 2009). These are:

- positive interdependence
- individual accountability
- equal participation (each pupil is expected to contribute)
- simultaneous interaction (as many pupils as possible should be actively involved at any one time)

The importance of social skills training and group processing in also highlighted in the Kagan approach.

The Kagan structures are designed to promote the four PIES principles through their very nature, and it is the presence of these four principles which indicate that CL is occurring (Kagan and Kagan, 2009).

Examples of Kagan structures include ‘Numbered Heads Together’, ‘Rally Robin’, and ‘Timed Pair Share’ (see Appendix K). In order to further illustrate the ways in which the use of Kagan structures themselves may facilitate CL and the presence of the PIES principles, the following scenario is given:

In a classroom situation, a class teacher may ask children to briefly discuss a question with a talk partner, for example, pupils may be asked to work with their partner to verbally list some multiples of 5 in maths, to recap examples of liquids in science, or to spot examples of verbs in a text during a literacy lesson.

When working with their talk partner, pupil A may dominate the conversation, and provide most, if not all, of the answers. If pupil A singlehandedly lists numerous examples on his/her own whilst pupil B listens, they may well have achieved a degree of superficial success in the academic aspects of the task. However, this would be viewed by Kagan and Kagan (2009) as small group, paired work rather than CL because there is no positive interdependence, that is, neither pupil A nor pupil B *need* the other to complete the task. Similarly, there is no individual accountability, meaning that pupil B is able to sit back in a ‘hitchhiking’ and ‘social loafing’ position. Neither equal participation nor simultaneous interaction are present, as only pupil A is actively contributing while pupil B takes a more passive role.

This appears in line with social interdependence theory, which suggests that with an absence of goal interdependence (success for pupil A is unrelated to that of pupil B), there will be individualistic effort and limited interaction (Johnson & Johnson, 2009).

Yet, this scenario can be transformed into a CL activity through the use of a simple Kagan structure, such as ‘Rally Robin’. In this structure, pupils may again work with their talk partner to list or recap information, but the Kagan structure means that they are expected to take it in turns to contribute, as in a tennis rally. Pupil A provides one answer, and then pupil B needs to provide an answer, before it is pupil A’s turn once more. If needed, pupil A may support his/her partner through providing hints in a ‘tip tip tell’ process, but there is a

clear requirement and expectation that pupil B will be active in the conversation and provide ongoing responses. This means that pupil A now needs pupil B in order to succeed, and vice versa (positive interdependence), and pupil B has a stake in the process and final outcome too (individual accountability). Pupil B is no longer able to sit back and let pupil A do all the work, as without pupil B's participation, the pair can only list the one initial answer given by pupil A – and therefore cannot achieve success in the academic aspects of the task. Similarly, the use of the 'Rally Robin' structure means that both pupils are now participating equally (equal participation) and simultaneous interaction is maximised.

In this CL scenario, through the use of the Kagan 'Rally Robin' structure, positive interdependence is created meaning that pupil A and pupil B perceive that achieving their goals is inextricably linked to their partner achieving theirs. Thus, in line with social interdependence theory, there is cooperative effort, with enhanced motivation to help each other to reach shared goals (Johnson & Johnson, 2009).

1.6.3.3 Slavin's Student Team Learning Approach and Curricular Approach

As well as the idea that pupils work together with responsibility for their own and their team mates' learning, Slavin's student team learning approach emphasises the use of team goals and team success with rewards to create and maintain pupils' motivation to take each other's achievement seriously (Slavin, 1995).

The team learning approach prioritises three elements:

- team rewards
- individual accountability
- equal opportunity for success

One example of a student team learning approach is student teams-achievement divisions (STAD). Over three to five lessons, pupils work cooperatively to ensure that they have all mastered the taught lesson content. They are then individually quizzed on the material, with rewards allocated to teams who have reached certain levels of performance.

Slavin's curricular approach includes specific approaches which can be used in maths, reading, and writing lessons.

The addition of a competitive element into the CL programme through team rewards is not one shared by all cooperative approaches, with Johnson & Johnson (1999) arguing that the use of rewards may negatively impact on future learning through a shift from intrinsic to extrinsic motivation.

1.6.3.4 A Summary of Cooperative Learning Models

The CL approaches described above are perhaps three of the most widely discussed (LaFont, Proeres, & Vallet, 2007). Within this, the bulk of literature and research concerns the Johnson and Johnson learning together approach, and Slavin's student team learning approaches (Murphy, Grey, & Honan, 2005; Slavin, 1995).

Despite some differences in approaches, particularly around the role of rewards and the degree of structure provided, there is general agreement that positive interdependence and individual accountability are particularly vital elements in effective CL (Jolliffe, 2007).

1.6.4 Implementation Considerations

In addition to the key elements and different models of CL discussed earlier, there are two further key areas of note in CL literature. These are the structuring of groups, and potential difficulties in implementation. These will now be briefly discussed.

1.6.4.1 Group Composition

CL approaches promote the use of small, teacher selected groups that are heterogenous in composition and maximise diversity (Reilly & Mitchell, 2010). Indeed, literature indicates that when choosing their own groups, pupils tend to choose those whom they know and feel are like themselves, leading to homogeneous groups which promote or reinforce status hierarchies (Kagan & Kagan, 1994). A meta-analysis of 'within-class grouping' has pointed to larger effects when group formation is on more than ability alone, highlighting the importance of teacher judgements of compatibility in promoting positive group interaction (Lou, et al., 1996).

There is evidence that pupils work best and achieve more when they are in groups of three to four. In a meta-analysis, Lou, et al. (1996) concluded that, on average, pupils placed in smaller groups achieved more, held more positive attitudes, and reported higher general self-concept than pupils in non-grouped or larger grouped situations.

1.6.4.2 Potential Difficulties in Implementation

Even with cooperative structures, there is no guarantee that pupils will engage in positive interactions (Slavin, 1983), and there always exists a risk of negative interaction which could lead to decreases in peer acceptance and liking, and decreases in learning and achievement (Battistich, Solomon, & Delucchi, 1993).

As with any classroom intervention, the teacher is critical to the effective implementation of CL (Battistich, Solomon, & Delucchi, 1993).

Yet the potential difficulties in implementing peer based pedagogies within a classroom context are acknowledged in the literature, with Cohen (1994) noting that it is a challenge that many teachers find difficult to accomplish. Indeed, Graves & Graves (1990) argue that even for 'good' teachers, it will take two

years before cooperative methods are ingrained in their practice, and Johnson, Johnson, & Holubec (1994) point towards several years of training and support to achieve expertise and mastery.

Antil, Jenkins, Wayne, & Vadasy (1998) caution that with more abstract models of CL, teachers may modify approaches in ways which detract from the essential principles and processes, and therefore reduce any effectiveness of the approach. In their study, they found that although teachers said that they used CL daily, further investigation revealed that it was not being implemented with fidelity nor to any recognised model.

Nevertheless, many studies have shown social and academic improvements over relatively short periods of time, with Rzoska & Ward (1991) pointing towards studies which have reported significant changes in as little as six one-hour sessions, or over only a three-week intervention period. Indeed, Jordan & Le Metais (1997) argue that effective strategies *can* be put into place over a short period of time, and gradually introduced in to normal classroom practice, given an appropriate framework.

1.6.5 Reported Impacts of Cooperative Learning

Proponents of CL are clear that it is an instructional strategy which promotes learning and development across a range of academic and social domains (Murphy, Grey, & Honan, 2005), with Johnson & Johnson (2000) arguing that no other pedagogical practice simultaneously achieves such diverse outcomes.

Gillies (2016) points to the meta analyses carried out by Johnson, Maruyama et al. (1981), Johnson & Johnson (2002), Roseth, Johnson, & Johnson (2008), and Slavin (1989) in concluding that there are a range of benefits for pupils in cooperating with others, and notes that there is overwhelming evidence that CL can have a profound effect on student learning and socialisation (Slavin, 2014).

Impacts from both academic and social domains will now be discussed, before considering whether CL differs in its impact for different groups of pupils.

1.6.5.1 Academic Outcomes

In a meta-analysis of 158 studies looking at academic achievement since 1970, Johnson & Johnson (2002) reported that all of the studies included in the review had significant positive impacts on student achievement, concluding that this gave strong validation for the impact of CL.

Despite naming Kagan CL structures as a main CL approach, Johnson & Johnson's (2002) meta-analysis did not find any studies evaluating the impact of this approach. They warned that although current research suggests that the likelihood of positive results following CL intervention is quite high, results are not guaranteed, and there is no reason to expect that all operationalisations of CL will be effective, recommending that the unevaluated CL methods, therefore, should be used with some caution.

Evaluation studies focusing on academic outcomes following CL have reported that, in comparison to their peers in control or individualistic conditions, pupils in CL conditions had:

- higher recall and higher reasoning scores (Johnson, Johnson, & Taylor, 1993)
- greater higher-level explanation and analytical thinking, and greater factual retention (Peterson, Johnson, & Johnson, 1991)
- greater increases in maths achievement (Jacobs, Watson, Sutton, & Jones, 1996)
- greater academic progress in science (Baines, Blatchford, & Chowne, 2007)

There is some evidence in the literature that gains made may endure over time, with a two-year longitudinal study by Thurston et al. (2010) reporting that social, cognitive, and affective gains made by CL groups in primary school remained evident once they had moved to secondary.

1.6.5.2 Social Outcomes

Johnson, Johnson, & Stanne (2000) argue that CL is crucial to preventing and alleviating a range of social difficulties encountered by children and young people.

Johnson & Johnson (2005) argue that there is considerable evidence that more academic and personal social support is experienced from peers and teachers in cooperative than in individualistic and competitive situations, with cooperative relationships ‘a prerequisite to creating a meaningful social support system for each student’ (Johnson & Johnson, 1999, p.64).

This is in line with work by Battistich, Solomon, Watson, & Schaps (1997) who note that teacher practices are able to impact directly on pupils’ sense of belonging, with classroom level practices which promote warmth and supportiveness, prosocial values, cooperation, and expression of ideas, crucial.

Johnson, Johnson, & Holubec (1994) argue that caring, committed, and supportive personal relationships grow from a sense of mutual accomplishment, mutual pride in joint work, and the bonding that results from cooperative effort, meaning that the more often students work in cooperative groups, the more they will like each other. Through cooperation, and the increased communication that this brings, pupils develop more multidimensional views of their peers, and feel increased acceptance – with enhanced expectations of future rewarding interactions.

Evaluation studies focusing on social outcomes following CL have reported that, in comparison to their peers in control or individualistic conditions, pupils in CL conditions had:

- a greater sense of cohesion (Johnson, Johnson, & Taylor, 1993)
- greater self esteem (Lazarowitz, Hertz-Lazarowtiz, & Hugh, 1994)

- greater peer acceptance and number of friendships (Lazarowitz, Hertz-Lazarowtiz, & Hugh, 1994)
- greater perceptions of help, encouragement, care, and affection from peers (Johnson, Johnson, & Anderson, 1983)
- more varied social interactions, and a greater willingness to engage with pupils outside of their immediate friendship group (Jordan & Le Metais, 1997)
- greater perceptions of social support (Johnson, Johnson, Buckman, & Richards, 1985), which increased with increasing frequency and longevity of CL

Other studies have reported some positive social outcomes in some classes in some domains, but not in others. Comparing gains between cooperative and control classes, Jacobs et al. (1996) reported mixed results, with at least one of the three intervention classes having significantly greater increases in self-concept, in attitudes towards maths, or in friendships, but with no class making statistically significantly different gains on all three outcomes.

Similarly, although Lazarowitz, Hertz-Lazarowtiz, & Hugh (1994) reported greater peer acceptance and friendships following CL, they found no significant differences between pupils in CL or individualistic conditions regarding perceptions of class cohesion.

Teacher – pupil and pupil – pupil interactions in CL groups

Hertz-Lazarowitz & Shachar (1990) found that when pupils worked in CL groups, teachers became 'facilitators of learning' rather than 'dispensers of knowledge', with increased pro social and positive instructional behaviours, communicating more positive, personal, and supportive messages to their pupils than when they were teaching whole class lessons. Changes in teacher practice, facilitated by the use of CL, may be important in peer acceptance, with instructional strategies communicating the notion that all pupils have equal status (Cohen, 1994), and with rules and structures in CL which mean that inclusion is encouraged and maximised, helping to remove hierarchies in a peer

group, and thus increase both participation and peer acceptance for low status pupils (Harrist & Bradley, 2003).

Furthermore, literature suggests that pupils in CL classrooms become more personally active in their dialogic exchanges (Hertz-Lazarowitz & Shachar, 1990). Beyond the academic benefits of more frequent task related interactions highlighted by Cohen (1994), Alexander (2008) argues that increased communication and peer talk – something which is promoted through CL structures – leads to a sense of solidarity and cohesion developing between pupils (Gillies, 2015), supporting pupils in developing, establishing, and sustaining relationships.

When pupils are working in a supportive environment where they know their opinions and contributions will be respected and valued, Gillies (2015) argues that dialogue increases, with increasingly multi-dimensional, dynamic, and vibrant interactions. A study by Gillies (2006) found that, in comparison to pupils in group work only conditions, pupil to pupil interactions were greater in CL groups, with dialogue characterised by three times as many elaborations, twice as many short responses, increased engagement behaviours, and more verbal behaviours which were regarded as helpful and supportive to peers and their group.

A dialogic approach to teaching and learning where active pupil participation is maximised (Skidmore & Gallagher, 2005), and pupils are actively engaged in reflection, discussion, dialogue and argumentation (Mercer & Littleton, 2007) is facilitated in classrooms where there is an ethos of collective, reciprocal, supportive and structured exchanges which are respectful and purposeful (Alexander, 2008). That is, the classroom climate needs to enable pupils to address learning tasks together, listen to each other, share ideas freely without worrying about criticism or negative judgement, and reach common understandings – aspects which CL may be seen as potentially able to support (Johnson & Johnson, 2013).

However, it is important to acknowledge that whilst CL is heavily dependent upon pupil interactions where pupils perceive equal status of all members of the group, this does not necessarily always occur (Jules, 1991). Indeed, although CL may be viewed as supporting equality of status, studies suggest that significant others, including teachers, can also strongly impact on peer ecology and peer acceptance. Studies have indicated that the behaviour of a teacher towards a pupil or pupils can set a model for peers to follow, either increasing or decreasing the likelihood of a child become more accepted by peers over time (Hughes, Cavell, & Willson, 2001), and the more differentially teachers behave, the more hierarchical the peer ecology is (Hendrickx, Mainhard, Boor-Klip, Cillessen, & Brekelmans, 2016).

1.6.5.3 Differential Impact

Many studies have investigated whether there is a differential impact of CL depending on pupils' initial academic ability. The literature is mixed, with Shachar (2003) looking at eight Israeli studies and concluding that the low achieving consistently emerged as those who derived maximum benefit from CL, with high achievers not receiving any significant benefit when compared to whole class instruction. Similarly, Lou, et al.'s (1996) meta-analysis concluded that whilst low-ability pupils learned more in heterogeneous ability groups, medium-ability pupils benefited significantly more in homogeneous groups.

However, this picture is contentious, with other studies reporting learning gains to all pupils in heterogeneous ability groups, with medium ability pupils just as active as their high ability peers (Gillies & Ashman, 1997). Indeed, Webb & Cullican (1983) noted that pupils in homogenous ability groups received no response to questions four times more often than their peers in heterogeneous ability groups. More recently, in a review of research, Gillies (2016) concluded that medium-ability students appear to do equally well in mixed- or same-ability groups, with CL benefitting all groups of pupils (Kagan & Kagan, 2009).

The relationship between pupils' social status and the impact of CL appears to have been less studied, but there are some indications of potential differences.

A study by Jules (1991) found an overall decrease in negative perceptions of peers and a greater interactive network in CL intervention classrooms, but within this, specific children remained very popular, or, as more was the case, very unpopular. That is, peer acceptance levels of the children did not all appear to increase in the same way, with the least accepted pupils remaining particularly stable in their status.

An explanation offered for this finding was that, in all cases, the least accepted children did not perceive themselves as having equal status in the group, and this may have impacted on the interaction dynamics. Indeed, Johnson & Johnson (1999) highlight the importance of positive goal interdependence and situational status equality as essential conditions for changing attitudes and acceptance of rejected children. Without this, Erwin (1998) cautions that the greater social freedom granted by CL may actually result in more children becoming socially isolated, with status differences intensifying the longer the group exists (Peterson, Johnson, & Johnson, 1991).

Further examples come from Andre, Deneuve, & Louvet (2011) who point to a series of studies carried out by Garel (2001) where CL approaches to integrating pupils with learning disabilities in to mainstream PE lessons did not always lead to better acceptance. On the contrary, it sometimes caused the pupils to be more rejected by their peers.

1.7 Conclusions: A Joining of Cooperative Learning with Peer Acceptance and Peer Support

The literature thus far has described peer acceptance and peer support, and their associations with a range of academic, social, and psychological outcomes. The ways in which peer acceptance and peer support may, or may

not, develop has been discussed through the use of a model of social exchange (Dodge et al.,1986).

There are similarities which can be drawn between the development of peer acceptance, and the access to and perceptions of peer support, with them both being dependent not only on the individual characteristics of a child, but also intricately linked with the peer group around them.

An outline of CL, including its background, theoretical perspectives and main approaches has been given, along with research regarding the potential impacts of CL in both academic and social domains.

The CL literature points to a potentially powerful stage upon which pupils can practice and perfect their social skills (Jules, 1991), with the ability to address key aspects of the Dodge et al. (1986) model in supporting the development of positive peer relationships, including acceptance and feelings of support, through carefully structured group work based on positive interdependence and individual accountability.

In order to further understand the evidence base regarding the impact of CL on peer acceptance and perceptions of peer support, a systematic literature review was conducted.

1.8 Systematic Literature Review

A systematic literature review allows researchers to conduct a transparent and replicable search of existing research, as related to a specific research question (Petticrew & Roberts, 2006). Through explicit and rigorous methods, including clear inclusion criteria, a systematic review aims to locate, appraise, and synthesise the best available evidence related to a specific question (Boland, Cherry, & Dickson, 2014). As well as identifying methodological strengths and weaknesses, systematic comparison of studies allows researchers to identify

gaps in knowledge and research, draw evidence based conclusions, and highlight areas for future work (Gough, Oliver, & Thomas, 2013).

The aim of the current systematic review was to systematically identify and synthesise research which looked at the impact of CL on peer acceptance and perceptions of peer support.

1.8.1 Review Question

What is the impact of CL on pupils' peer acceptance and perceptions of peer support?

1.8.2 The Search Strategy

A final search for studies relevant to the research question was carried out on 29th January 2017. The following electronic databases were searched:

- PsychINFO
- Web of Science
- ERIC (EBSCO)

The following search terms were used:

‘cooperative’ OR ‘co-operative’ (title)

AND

‘learning’ (title)

COMBINED WITH *AND*

‘support*’ OR ‘accept*’ OR ‘relation*’ OR ‘friend*’ OR ‘cohesi*’ OR ‘connect*’(keyword / anywhere)

The search returned 1341 records.

The reference lists of the full text articles retrieved were scanned. This highlighted three further papers.

1.8.3 Selection Criteria

Inclusion criteria were developed with regard to the review question, as follows:

Primary inclusion criteria:

1. The study was published in a peer reviewed paper
2. The full paper was available in English
3. The study was published between 1990 and January 29th 2017

Secondary inclusion criteria:

4. The study evaluated the impact of a CL intervention with regards to review relevant outcomes
5. The study employed an experimental design, with repeated measures
6. The intervention took place in a school setting with a school based sample, and with participants aged between four and sixteen years old

Inclusion criteria were developed in order to increase the likelihood that the studies would be of relatively high quality, and allow an evaluation of the impact of CL according to the review question, in a population of primary interest and relevance to an EP.

A flow diagram of the study selection process is contained in Appendix A.

The full text of 29 papers which met the inclusion criteria were obtained. Of these, 16 papers were excluded for reasons such as only using post intervention measures, for outcomes which were not related to the review question, using participants over the age of 16, or not evaluating a CL intervention.

References of studies excluded at full text are contained in Appendix B.

1.8.4 Studies Included in the Systematic Literature Review

Thirteen studies met the inclusion criteria and were included in the systematic review. Authors, titles, and overall weight of evidence are shown in Table 1.1.

For detailed summary data for each of the thirteen studies, please refer to Appendix C (sample, design, and intervention) and Appendix D (measures, DVs, and relevant findings).

Table 1.1 Author, title, and overall weight of evidence for studies included in the systematic review

Author	Title	Overall WoE D
Andre, Deneuve, & Louvet (2011)	Cooperative learning in physical education and acceptance of students with learning disabilities	High
Andre, Louvet, & Deneuve (2013)	Cooperative group, risk taking and inclusion of pupils with learning disabilities in physical education	Medium
Bertucci, Conte, Johnson, & Johnson (2010)	The impact of size of cooperative group on achievement, social support, and self-esteem	High
Bertucci, Johnson, Johnson, & Conte (2011)	The effects of task and resource interdependence on achievement and social support: An exploratory study of Italian children	High
Bertucci, Johnson, Johnson and Conte (2012)	Influence of group processing on achievement and perception of social and academic support in elementary inexperienced cooperative learning groups	Medium

Jacobs, Watson, Sutton, & Jones (1996)	Effects of a cooperative learning method on mathematics achievement and affective outcomes of students in a private elementary school.	High
Jacques, Wilton, & Townsend (1998)	Cooperative learning and social acceptance of children with mild intellectual disability	High
LaFont, Proeres, & Vallet (2007)	Cooperative group learning in a team game: Role of verbal exchanges among peers	Medium
Oortwijn, Boekaerts, Vedder, & Fortuin (2008)	The impact of a cooperative learning experience on pupils' popularity, non-cooperativeness, and interethnic bias in multiethnic elementary schools.	Medium
Piercy, Wilton, & Townsend (2002)	Promoting the social acceptance of young children with moderate–severe intellectual disabilities using cooperative-learning techniques	High
Putnam, Markovchick, Johnson, & Johnson (1996)	Cooperative learning and peer acceptance of students with learning disabilities	High
Topping et al. (2011)	Cooperative learning in science: Intervention in the secondary school	High
Vasileiadou (2009)	Cooperative learning and its effects on pre-primary marginalized children	Low

1.8.5 Quality and Relevance of Included Studies

Harden & Gough (2012) note that 'assessing the quality and relevance of individual studies within a review is a crucial part of the review process' (p.154).

The Gough (2007) WoE framework was used to assess the methodological quality, appropriateness, and relevance of each paper to the review question. The WoE strands are summarised in Table 1.2.

Table 1.2 A description of each Weight of Evidence strand (A to D)

Weight of Evidence Strand	Description
WoE A: Methodological Quality	Generic judgement about the coherence, quality, and integrity of the study
WoE B: Methodological Appropriateness	Appropriateness of the study design for answering the review question.
WoE C: Relevance of Evidence	Relevance of the study in answering the review question.
WoE D: Overall Weight of Evidence	Overall assessment of the study based upon the ratings given for A, B, and C.

Criteria developed by the researcher in order to make WoE judgements can be found in Appendix E. This was informed by Gersten et al.'s (2005) quality indicators for group experimental and quasi-experimental research, and by Downs & Black's (1998) checklist of methodological quality for randomised and non-randomised studies.

For information as to how these judgements were made for each individual study, including ratings on each WoE strand, please refer to Appendix F.

With the exception of Vasileiadou (2009) which was rated low in all strands, and Andre et al. (2013) which was judged as having low relevance to the review question (WoE C), all studies received medium or high judgements across the strands.

The systematic review adopted a weighting approach to synthesis whereby all thirteen studies contributed, but with less weight given to the study assessed as low WoE D (Vasileiadou, 2009).

1.8.6 Synthesis of Results

In order to synthesise the research evidence, the thirteen studies were compared on the following themes:

- sample
- design
- intervention
- measures and DVs
- relevant findings

The synthesis process aimed to bring clarity regarding the overall evidence base, as relevant to the review question, and in doing so, inform the research questions and design of the current study.

A narrative summary of pertinent points is now presented. As only three of the studies looked at perceptions of peer support, these have been summarised distinctly at the end of each subsection.

1.8.6.1 Sample

Peer Acceptance:

Only one study was carried out in the UK, and this was based in Scotland (Topping et al., 2011). The other studies were carried out in France (Andre et al., 2011; Andre et al., 2013; LaFont et al., 2007), New Zealand (Jacques et al., 1998; Piercy et al., 2002), the USA (Jacobs et al., 1996; Putnam et al., 1996), the Netherlands (Oortwijn et al., 2008), and Cyprus (Vasileiadou, 2009).

Of the six studies carried out in the last ten years (Andre et al., 2011; Andre et al., 2013; LaFont et al., 2007; Oortwijn et al., 2008; Topping et al., 2011; Vasileiadou, 2009), three were in France, and involved non-academic lessons (Andre et al., 2011; Andre et al., 2013; LaFont et al., 2007), and one was rated as low WoE across all strands (Vasileiadou, 2009).

Participants came from a variety schools, including pre-primary/kindergarten, primary, middle, or elementary schools, and secondary schools. The age of participants ranged from 5 years old (Piercy et al., 2002; Vasileiadou, 2009) to 10 and 11 years old and above (Andre et al., 2011; Andre et al., 2013; Oortwijn et al., 2008; Putnam et al.,1996; Topping et al., 2011). Three studies had pupils aged between 8 and 11 years old (Jacobs et al., 1996; Jacques et al., 1998; LaFont et al., 2007).

The number of participants ranged from 13 (Vasileiadou, 2009) to 644 (Topping et al., 2011). Five studies involved more than 150 participants (Andre et al., 2011; Andre et al., 2013; Jacobs et al., 1996; Putnam, et al.,1996; Topping et al., 2011), and five studies ran in multiple schools (Andre et al., 2011; Jacques et al., 1998; Oortwijn et al., 2008; Putnam et al.,1996; Topping et al., 2011).

In the majority of studies, there was an approximate balance between male and female participants, although two studies had entirely (Piercy et al., 2002) or predominantly (Andre et al., 2013) male participants.

Perceptions of peer support:

Two studies were carried out in Italian secondary schools, and one in an Italian elementary school (Bertucci et al., 2012). The studies ran in individual schools, with 62 (Bertucci et al., 2010), 66 (Bertucci et al., 2011), and 61 (Bertucci et al., 2012) participants in each study, with an approximate balance between males and females.

1.8.6.2 Design

Peer Acceptance:

All studies had pre and post-test measures, with either random assignment to condition or a quasi-experimental design. One study included a follow up at five weeks (Jacques et al., 1998).

Seven of the ten studies had control or comparison groups, with Piercy et al. (2002) having both a control and 'social contact' comparison group. In the remaining three studies, all participants received CL (Andre et al., 2013, Oortwijn et al., 2008; Vasileiadou, 2009).

Four of the studies used intact class groupings, and were taught by their regular teacher (Jacobs et al., 1996; Oortwijn et al., 2008; Putnam et al.,1996; Topping et al., 2011). Due to their designs, other studies did not use intact classes. In three studies, special population pupils (e.g. SEGPA, and 'learning disabled') moved in to the usual intact classes for the research period (Andre et al., 2011; Andre et al., 2013; Jacques et al., 1998). SEGPA pupils are those pupils in French middle schools who have been placed in special classes based on their writing and reading levels (Section d'Enseignement Générale et Professionnelle Adaptee). In Piercy et al. (2002), mainstream children went in to the special education class to carry out practical group work. In the remaining studies, participants were taken from a range of different classes and put in to new class groups or teams for the purposes of the research.

Perceptions of peer support:

All three studies which had perceptions of peer support as their DV used a pre-test post-test design. Two employed stratified randomised assignment to condition (Bertucci et al., 2010; Bertucci et al., 2011), and Bertucci et al. (2012) used randomised assignment. Pupils were not in their usual class groupings in any study. Two studies had two intervention groups as well as a control group: Bertucci et al. (2010) compared CL in pairs to CL in groups of four; and Bertucci et al. (2011) compared task and goal interdependence with resource and goal interdependence. In these two studies, pupils were described as mainstream pupils, of 'normal intelligence or above'. Bertucci et al. (2012) had two intervention conditions and no control, comparing CL with and without group processing.

1.8.6.3 Intervention

Curriculum Focus: In half of the peer acceptance studies, CL interventions were carried out during academic lessons (Jacobs et al., 1996; Jacques et al., 1998; Oortwijn et al., 2008; Putnam et al.,1996; Topping et al., 2011). In the three French studies, the interventions were carried out during PE lessons (Andre et al., 2011; Andre et al., 2013; LaFont et al., 2007). In other studies, specially devised activities were planned for participants to engage in, for example, drawing and painting (Piercy et al., 2002; Vasileiadou, 2009).

In two of the studies looking at perceptions of peer support, pupils were taught a unit about alcohol, tobacco, and drug abuse (Bertucci et al., 2010; Bertucci et al., 2011). In Bertucci et al. (2012), CL was used within usual academic lessons.

Training and Facilitators: In the peer acceptance studies, the amount of training that staff received to support them in implementing CL varied greatly. It ranged from a two-hour training session (Oortwijn et al., 2008) up to 34 hours over a six month period (Putnam et al.,1996). Other studies included a number of days training, and ongoing weekly support.

In most of the studies, class teachers delivered the CL sessions. The only exceptions to this were Jacques, Wilton, & Townsend (1998) and Vasileiadou (2009) where the intervention was delivered by the researcher, and staff were not trained in CL.

Of the seven studies which had more than one condition, in four studies, class teachers taught in all conditions (Andre et al., 2011; LaFont et al., 2007; Piercy et al., 2002; Putnam et al.,1996).

In two of the studies which looked at perceptions of peer support, teachers delivered the lessons but no information was given regarding training (Bertucci et al., 2010; Bertucci et al., 2011). In Bertucci et al. (2012), the intervention was researcher led.

Intervention Model: Six studies of the peer acceptance studies reported that the CL intervention was based on Johnson and Johnson's five principles (Andre et al., 2011; Andre et al., 2013; LaFont et al., 2007; Piercy et al., 2002; Putnam et al.,1996; Vasileiadou, 2009). Two studies based their CL intervention around Slavin's STAD method and his jigsaw activities (Jacobs et al., 1996; Jacques et al., 1998). The remaining two studies did not clarify the approach taken to CL (Oortwijn et al., 2008; Topping et al., 2011).

The three perceptions of peer support studies were based on the Johnson and Johnson model, with some information given as to how goal, task and/or resource interdependence was created (Bertucci et al., 2010; Bertucci et al., 2011; Bertucci et al., 2012).

Social Skills Training: Only three of the peer acceptance studies made reference to initial social skills training. This ranged from two lessons (Oortwijn et al., 2008), to one week (Vasileiadou, 2009), or one to two months (Topping et al., 2011). Piercy et al. (2002) clearly note that there was no social skills training involved, and there was no mention of social skills training in the remaining studies.

There was no mention of social skills training in any of the perceptions of peer support studies (Bertucci et al., 2010; Bertucci et al., 2011; Bertucci et al., 2012).

Group Size and Composition: Seven of the ten peer acceptance studies placed pupils in groups of three to four (Andre et al., 2011; Andre et al., 2013; Oortwijn et al., 2008; Piercy et al., 2002; Putnam et al., 1996; Topping et al., 2011, Vasileiadou, 2009), and the others in groups of four to six (Jacobs et al., 1996; Jacques et al., 1998; LaFont et al., 2007).

Six studies described teacher selection of groups to ensure heterogeneity, primarily based upon relevant ability (Jacques et al., 1998; LaFont, et al., 2007; Piercy et al., 2002; Vasileiadou, 2009), with two of these studies also considering gender (Jacobs et al.,1996) and friendships (Topping et al., 2011). In three further studies, groups included either a 'SEGPA' pupil or a pupil who was experiencing learning difficulties, thus providing a degree of heterogeneity of ability (Andre et al., 2011; Andre et al., 2013; Putnam et al.,1996).

With regards to perceptions of peer support, Bertucci et al. (2010) varied the group size, with pupils working either in pairs or in fours. Bertucci et al. (2011) did not clarify group size. Both studies described stratified random assignment to condition, but it was not clear if this also included to group. Pupils in Bertucci et al. (2012) worked in groups of three to four, with pupils randomly assigned to a CL work group.

Frequency and Duration: Intervention duration in the peer acceptance studies varied from three active intervention weeks (Vasileiadou, 2009), up to five to six months (Topping et al., 2011), and eight months (Putnam et al.,1996). The remaining studies lasted between five and ten weeks, with the possible exception of Oortwijn et al. (2008) where no information was given regarding the duration of intervention.

In the six studies where there was sufficient information to identify intervention frequency, this varied from daily CL sessions (Vasileiadou, 2009), to studies where sessions were once a week (LaFont et al., 2007), twice a week (Piercy et al., 2002; Putnam et al.,1996), and four times a week (Jacobs et al., 1996; Jacques et al., 1998).

Individual session duration was provided for all peer acceptance studies, with the exception of Topping et al. (2011). Session duration varied from 20 minutes (Jacobs et al., 1996) up to 2 hours (Andre et al., 2011; Andre et al., 2013).

There was enough information to calculate approximate exposure times for the peer acceptance studies, with the exception of Topping et al. (2011). Approximate exposure times varied from six hours (Vasileiadou, 2009) up to 42-48 hours (Putnam et al., 1996). Five of the nine studies had an approximate exposure time of between 12 and 14 hours (Andre et al., 2011; Andre et al., 2013; Jacobs et al., 1996; Jacques et al., 1998; Piercy et al., 2002).

With regards to perceptions of peer support, two studies lasted six weeks, with one intervention session a week. It is likely that sessions lasted 90 minutes in each study, although there is an inconsistency in the text of Bertucci et al. (2011) suggesting that sessions may have been 45 minutes. Therefore, approximate exposure was 9 hours in Bertucci et al. (2010), and between 4 ½ and 9 hours in Bertucci et al. (2011). In Bertucci et al. (2012), pupils received five 90 minute sessions over a three week period, with approximately 7 ½ hours exposure.

Intervention Fidelity: The use of intervention fidelity checks was noted in four peer acceptance studies. In two of these studies, observations confirmed that treatment conditions were met (Andre et al., 2011; Oortwijn et al., 2008). In the other two studies, observations were carried out, but the outcomes and conclusions of these not given (Putnam et al., 1996; Topping et al., 2011).

With regards to perceptions of peer support, two of the three studies noted fidelity checks but neither indicated the outcomes or conclusions of these (Bertucci et al., 2010; Bertucci et al., 2011).

1.8.6.4 Measures and Dependent Variables

All studies included in the review collected quantitative data. In Vasileiadou (2009), this was supplemented through observations which were used to triangulate data, and in LaFont et al. (2007), quantitative data was supplemented through the use of sociograms.

With regards to peer acceptance, three studies cited research and literature which supported the reliability and stability of the approach taken to measuring the DV (Jacques et al., 1998; Oortwijn et al., 2008; Topping et al., 2011). While other studies provided no specific information regarding reliability, their approaches are commonly used within peer acceptance research. The large majority of studies used forced choice nominations (Andre et al., 2011; Andre et al., 2013; LaFont et al., 2007; Oortwijn et al., 2008; Putnam et al., 1996; Topping et al., 2011; Vasileiadou, 2009), with other studies using limited peer nomination techniques (Jacques et al., 1998), a combination of techniques (Piercy et al., 2002), and an additional sociogram (LaFont et al., 2007). Jacobs et al. (1996) used a 'sociometric instrument' but provided no further details.

Two studies calculated the reliability of their measures, and concluded that their sociometric indices were stable (Andre et al., 2011; Andre et al., 2013). Both of these studies used the Frederickson and Furnham (1998) Social Inclusion Survey.

With regards to perceptions of peer support, all studies used the Classroom Life Instrument (Johnson, Johnson, & Anderson, 1983), reporting reliability indices of *a* = .67 academic, and *a* = .78 personal support (Bertucci et al., 2010; Bertucci et al., 2011; Bertucci et al., 2012).

1.8.6.5 Relevant Findings

With the exception of Vasileiadou (2009), all studies included in the review used inferential statistics to analyse their findings and establish statistical significance. These were through a range of correlations, t-tests, ANOVA and ANCOVA procedures.

None of the peer acceptance studies reported effect sizes, but all of the perceptions of peer support studies did (Bertucci et al., 2010, Bertucci et al., 2011; Bertucci et al., 2012). However, limited commentary was given on these.

Findings:

Ten studies looked at peer acceptance and related constructs such as friendship and popularity as a DV.

Three studies looked at perceptions of peer support as a DV.

Appendix D contains a key facts table summarising measures, DVs, and relevant findings for each study.

The narrative which follows provides a synthesis of study findings.

Peer Acceptance – Positive Impact of CL:

Eight studies reported positive impacts following CL on at least one of their review relevant DVs. Of these, six studies reported a statistically significant increase from pre to post-test on the:

- social preference scores of SEGPA students as rated by their mainstream peers in the intervention group. This was not found in the control group (Andre et al., 2011).

- social preference scores of SEGPA students as rated by their mainstream peers in the high risk CL condition. However, there was no significant increase for those in the low risk CL conditions (Andre et al., 2013).

- social acceptance of focus pupils with 'mild intellectual disability' by their peers in CL classes but not in control classes. This was maintained at 5 week follow up (Jacques et al., 1998).

- ratings of popularity, and a significant decrease in perceived non-cooperativeness in heterogeneous teams (Oortwijn et al., 2008).

- peer acceptance and popularity ratings for focus pupils with 'moderate to severe intellectual disability' by their mainstream peers, which was not seen in the control or comparison conditions (Piercy et al., 2002).

- work peer acceptance of mainstream and special education pupils when rated by mainstream peers in the intervention but not control group (Putnam et al.,1996).

A seventh study reported significantly greater gains in friendships of children in CL than control classes in Grade 4 (Jacobs et al., 1996). However, no significant differences were found in gains between the two conditions in Grades 3 or 5, and there was limited information on how friendship was measured.

An eighth study reported that nine out of the seventeen pupils received an increased number of peer acceptance nominations from pre to post-test (Vasileiadou, 2009). Of the four children identified as rejected, three received more votes from peers post-intervention. The fourth child had no change in votes, but observations indicated increased playground interactions. However, no statistical analyses were conducted, and as there was no control group any changes may have been due to extraneous factors. This contributed to the low WoE rating for this study.

It must be noted that two of the studies were concerned only with outcomes for specific populations, that is, pupils with 'mild intellectual disability' (Jacques et al., 1998), and pupils with 'moderate to severe intellectual disability' (Piercy et al., 2002). This limits the extent to which these results can be generalised to mainstream populations.

In another three of the six studies, although there was a focus on special populations, that is SEGPA students (Andre et al., 2011; Andre et al., 2013), and special education pupils classified as 'disabled' (Putnam et al.,1996), data was collected regarding the outcomes for mainstream peers too. However, nominations were only made by mainstream pupils which continues to pose a limitation to interpretations of overall data, and outcomes for mainstream peers were broadly non significant.

In the case of Andre et al. (2013), it is important to note that although a significant increase pre to post-test for social preference of SEGPA pupils in the high risk CL condition was reported, there was no significant change found in the low risk CL condition. In the high risk condition, pupils were encouraged to perform gymnastic routines at increasing height, whereas in the low risk condition, routines remained at the same height but became increasingly technical. Although this study concluded in support of CL, it was only under specific circumstances with specific pupils.

As Oortwijn et al. (2008) lacked a control group, extraneous variables and maturation effects could have confounded results.

Peer Acceptance - No Impact of CL:

Four studies reported no positive impacts following CL on at least one of their review relevant DVs (Andre et al., 2011; Andre et al., 2013; LaFont et al., 2007; Topping et al., 2011). These studies reported no statistically significant increases from pre to post-test on the:

- social preference or social impact scores of mainstream pupils (Andre et al., 2011; Andre et al., 2013).

- social impact scores of SEGPA students as rated by their mainstream peers (Andre et al., 2011; Andre et al., 2013).

- interpersonal relationships of pupils (LaFont et al., 2007).

- friendships of pupils, measured through the number of pupils they like working with in science, like spending time with at break, and like spending time with out of school (Topping et al., 2011).

Although Andre et al. (2011) did report positive impacts of CL on the social preference scores of SEGPA students as rated by their mainstream peers, these impacts did not extend to social preference scores of mainstream children in either this study nor in the study by Andre et al. (2013). Furthermore, neither study found any statistically significant change in the social impact scores of SEGPA students or mainstream children.

Although LaFont et al. (2007) reported no statistically significant increase in interpersonal relationships, as assessed by the number of attraction or rejection nominations, a closer look at the sociograms did indicate that there were increased attractions for two out of three intervention teams, and decreased attractions for two out of three control teams. The least successful intervention team decreased, and the most successful control team increased. Least skilful players in the control group remained or became isolated by post-test.

Topping et al. (2011) noted that any positive findings at an individual class level were almost equally balanced out by negative findings. Having found no statistically significant changes for intervention or control groups from pre to post-test, they then looked at differences between the two groups at post-test. This revealed that, at post-test, ‘like working with’ was significantly higher in the control group than intervention, whereas ‘break time’ and ‘out of school’ was significantly higher for the intervention than control group. This led to the

conclusion that CL appeared to enhance friendships beyond the work group only. However, analysis of post-test data only at this stage did not account for the impact of any pre-test differences between the groups.

Perceptions of Peer Support

Two of the three studies reported positive impact following CL on perceptions of peer support. These were as follows:

- Pupils who had worked in CL pairs and CL fours perceived significantly greater peer academic support than pupils who had been in the individualistic learning condition (Bertucci et al., 2010). This represented a large effect (Eta Squared = .22).

- Pupils in both CL conditions (resource and goal interdependence; task and goal interdependence) had greater perceptions of peer academic support than pupils in the individualistic learning condition (Bertucci et al., 2011). This represented a large effect (Eta Squared = .30).

Whilst the findings for peer academic support appeared consistent, the impact on peer personal support appeared less clear, as follows:

- no significant difference between CL and individualistic learning for peer personal support (Bertucci et al., 2010).

- pupils in the CL resource and goal interdependence had greater perceptions of peer personal support than those in the CL task and goal interdependence or individualistic learning conditions (Bertucci et al., 2011). This represented a large effect (Eta Squared = .19).

Bertucci et al. (2012) found no significant difference between CL with and without group processing for peer academic or peer personal support. With no control group, it is difficult to interpret this finding, with a lack of clarity as to

whether both conditions had increased by an equal amount or whether CL had had no impact on the DVs.

1.8.7 Strengths and Limitations of the Systematic Review

The systematic review used a clear and transparent strategy to locate studies, with explicit criteria when identifying studies suitable for inclusion. WoE was considered when interpreting studies.

However, despite best intentions, it is possible that the review was not exhaustive, with the possibility that some relevant studies may have been missed due to the search terms employed or the databases used.

Similarly, inclusion criteria may have meant that pertinent studies were overlooked, for example, because they were published prior to 1990, not written in English, or were unpublished. This may have changed the conclusions of the review.

It is also important to note that the author alone retrieved papers, assessed relevance, carried out data extraction, weighting, and synthesis. Reliability or validity of these decisions has not been verified.

1.8.8 Summary and Conclusions of the Systematic Review

In addressing the question of ‘what is the impact of CL on pupils’ peer acceptance and perceptions of peer support?’, the systematic search process identified thirteen studies suitable for inclusion in the review.

Summaries and conclusions are described for the peer acceptance literature, and then for the peer support literature.

Ten studies looked at peer acceptance or related constructs. Sociometric measures were used across the majority of studies, including both forced choice and limited peer nomination techniques.

The strengths of the literature with regards to CL and peer acceptance is that research has been carried out in a range of subjects (academic and non-academic), across a range of ages, and in a range of countries. Of the ten studies, nine found a positive effect of CL on at least one of their review relevant DVs – although four of these studies also reported no effect on an additional relevant DV.

However, it was particularly striking that only one study had been carried out in the UK (Topping et al., 2011). Moreover, of the six studies carried out in the last ten years, three of these were in France and involved physical education lessons, and a fourth had low WoE ratings across all strands.

Furthermore, only five studies were carried out in academic lessons, and, surprisingly only four were carried out in established, intact classes. Additionally, of the studies which reported positive impacts following CL, the majority related to the ratings given to pupils from special populations rather than the wider class.

The majority of CL interventions were based on the Johnson and Johnson model, with others noting the use of Slavin's STAD model. Yet, when discussing such conceptual approaches, Antil et al. (1998) note that without a great deal of training and support, teachers may modify the approach in such a way that they unintentionally remove some of the core processes and principles, reducing any impact. This may have had an impact on results, particularly given that some studies involved only a few hours training. However, at the same time, many studies involved an arguably unrealistic amount of input and ongoing support, with one study providing 34 hours input over six months. Difficulties in using the approaches may also explain why in two studies the researcher delivered the intervention.

With only four studies having reported conducting intervention fidelity checks, and only two reporting the outcomes of these, these is some uncertainty about what actually occurred during CL sessions.

It is also unclear how many studies included the pre-teaching of social skills as part of CL. It appears likely that only a minority did so, and, as literature indicates CL with social skills to be more efficacious, this may have impacted on findings (Gillies, 2016).

Methodological difficulties included studies without control or comparison groups where it is hard to establish if changes are different to that which may have naturally occurred, and studies where class teachers taught pupils in both conditions.

The systematic literature review has highlighted some potential for CL to impact on pupils' peer acceptance. However, studies do indicate some mixed findings, and the extent to which these can be generalised to a UK, mainstream population with intact classes, a nominal level of training, and teacher delivery is unclear.

Three studies looked at perceptions of peer support.

The strengths of the literature with regards to CL and perceptions of peer support is that the research used randomised designs with established measures. Two studies also used control groups.

However, there were only three studies identified, and they were all carried out in Italy by the same researchers, using the same measure. Two of the studies were particularly similar in that they used the same unit of work with 7th grade pupils described as being of average or above academic ability (Bertucci et al., 2010; Bertucci et al., 2011). Furthermore, in all three studies, the pupils were in classes arranged for the purposes of the research. These factors pose some limitations to the extent to which the findings of the three studies can be generalised.

Whilst indicating some potential for CL to impact on perceptions of peer support, particularly academic peer support, it must be the conclusion of the current review that the evidence base for the impact of CL on perceptions of peer support is tentative and in the early stages of development.

1.8.9 Unpublished Literature

Acknowledging a lack of UK based work regarding the impact of CL on peer acceptance and perceptions of peer support identified through the primary systematic literature review, a search of the 'ProQuest Dissertations and Theses: UK & Ireland' database was conducted (final search: 14th February 2017). This followed the same search strategy as the primary systematic search, with the exception of the need for a paper to be peer reviewed, and the inclusion of a UK context.

As unpublished literature has not been subject to scrutiny through a peer review process, it should be treated with appropriate caution. Nevertheless, such literature may offer an opportunity to correct for publication bias (Mertens, 2015), and, given the limited UK evidence base highlighted in the primary systematic review, may offer a wider understanding of the field.

This search revealed 22 results. Two of these met inclusion criteria and were relevant in addressing the review question.

Craig (2010) looked at the impact of CL on peer acceptance of Year 4 children in an English primary school. The study used a pre-test post-test non-equivalent groups quasi-experimental design, with peer acceptance measured using the SIS (Frederickson & Furnham, 1998), with supplementary measures using the Strengths and Difficulties Questionnaire looking at peer problems and pro-social behaviours. Pupils received a CL intervention based on Kagan approaches, with initial social skills training. Craig (2010) reported that pupils in the intervention group were significantly more accepted at post-test than pupils

in the control group, had significantly greater levels of pro-social behaviour, and significantly lower levels of peer problems.

Craig's (2010) study was limited with regards to external validity as it involved only two classes, with a total of 54 participants. With just one intervention teacher, Craig (2010) himself notes the possibility that general teacher characteristics and teaching style could have influenced outcomes, either exaggerating or underrepresenting any effects. Nevertheless, the results of the study are promising with regards to the impact of CL on peer acceptance in a UK, mainstream school context.

Kirk (1997) looked at the impact of CL on friendship patterns of 11-13 year old females at an Irish primary school. The study used a pre-test post-test control group quasi-experimental design, with friendship patterns and helpful working relationships assessed using sociograms. Pupils received a CL intervention based on an adaptation of the Johnson and Johnson learning together model which incorporated Slavin's reward structures. Kirk (1997) reported no major changes in friendship patterns, although there was a slight change in helpful working relationships. Structured observations carried out by the school principal suggested that pupils' awareness of social skills and attitudes towards work and others had increased.

Kirk's (1997) sample was limited as participants were all female, and there were only 16 pupils in each condition. Both experimental and condition groups were in the same, researcher taught classroom, with one group receiving CL for three weeks in maths, and the control group then receiving CL for three weeks in spelling sessions. Diffusion of treatment, as well as the short time scale may have played a part in the limited findings regarding the impact of CL.

1.9 Rationale for Current Research

The literature review has highlighted the importance of peer acceptance and perceptions of peer support in a range of areas.

The process of peer acceptance appears complex, with peer acceptance – or rejection - not explained purely in terms of a child's individual characteristics, but also as a social process within which peers' behaviours and perceptions are important influences (Bierman, 2004). Although many pupils are able to gain acceptance with their peers and access peer support systems, due to a potential myriad of reasons, other pupils are unable to do so (Johnson & Johnson, 1999).

With literature indicating that that both a child's peer group status and peer social support are moderately stable (Demaray, Malecki, Davidson et al., 2005; Herbert, 2002), there is a clear need to identify evidence based interventions which can facilitate and enhance peer acceptance and access to support for children experiencing difficulties in these areas.

CL has been highlighted as an intervention that could potentially support teachers in structuring the environment for successful peer interactions as well as providing pupils with the instruction and support they need to develop their social skills (Battistich & Watson, 2005). In doing so, CL appears to meet elements highlighted as important when intervening with peer acceptance and peer support (section 1.5), and has the potential to address multiple elements of the Dodge et al. (1986) model of social exchange. Furthermore, as it is a teaching approach which can be delivered by teachers to their intact classes, it has the potential to support these aspects of social development in a subtle and inclusive manner rather than through direct, individual focused, withdrawal based intervention (Piercy, Wilton, & Townsend, 2002).

However, whilst the systematic literature review indicates some promise for the impact of CL on children's peer acceptance, as described in section 1.8.8, there are weaknesses in the evidence base, particularly in the extent to which findings may be generalised to a UK, mainstream class, teacher led delivery model. There is some promising but scant research looking at the impact of CL on perceptions of peer support, and that which does exist is not UK based.

Furthermore, there are some indications of potential differences in impact for children with different initial peer acceptance levels. The impact of CL on the population with which an EP may typically work – children with low peer acceptance – is unclear, as are the relative impacts for their medium and high accepted peers. Given the importance of peer acceptance and perceived peer support, it is crucial that a well-meaning intervention does not exacerbate an already difficult situation for those most needing support, or lead to a widening gap between the least peer accepted children and their peers.

Therefore, the current research will consider the impact of CL on pupils' peer acceptance and perceptions of peer support. It will then consider the impact of CL on children with low initial levels of peer acceptance and perceptions of peer support, and consider whether there are any differences in the impact of CL for children with initially low, medium, and high levels of peer acceptance and perceptions of peer support.

1.10 Research Questions

These are outlined in Table 1.3.

Table 1.3 Research questions and hypotheses

Research Question	Hypotheses
Research Question 1: What is the impact of CL on pupils' perceptions of peer support?	**Experimental Hypothesis:** There will be a statistically significantly greater increase in pupils' perceptions of peer support following CL (experimental condition), when compared to pupils in a control condition. **Null Hypothesis:** There will be no statistically significant difference between changes in pupils' perceptions of peer support in the experimental or control conditions.
Research Question 2: Does the impact of CL on pupils' perceptions of peer support vary, depending on their initial levels of a) peer acceptance or b) perceptions of peer support?	2i **Experimental Hypothesis:** There will be statistically significant differences in the increases of pupils' perceptions of peer support, according to their initial levels of a) peer acceptance and b) perceptions of peer support. **Null Hypothesis:** There will be no statistically significant differences in the increases of pupils' perceptions of peer support, according to their initial levels of a) peer acceptance or b) perceptions of peer support.

	2ii **Experimental Hypothesis:** There will be statistically significantly greater increases in perceptions of peer support for pupils with low initial perceptions of peer support in the experimental than control group. **Null Hypothesis:** There will be no statistically significant differences in the increases in perceptions of peer support for pupils with low initial perceptions of peer support in experimental and control groups.
***Research question 3*:** What is the impact of CL on pupils' peer acceptance?	***Experimental hypothesis**:* There will be a statistically significantly greater increase in pupils' peer acceptance following CL (experimental condition), when compared to pupils in a control condition. ***Null hypothesis:*** There will be no statistically significant difference between changes in pupils' peer acceptance in the experimental or control conditions.
Research question 4: Does the impact of CL on pupils' peer acceptance vary, depending on pupils' initial levels of peer acceptance?	4i ***Experimental hypothesis**:* There will be statistically significant differences in the increases of pupils' peer acceptance, according to their initial levels of peer acceptance. ***Null hypothesis:*** There will be no statistically significant differences in the increases of pupils' peer acceptance, according to their initial levels of peer acceptance

	4ii **Experimental Hypothesis:** There will be statistically significantly greater increases in perceptions of peer acceptance for pupils with low initial levels of peer acceptance in the experimental than control group. **Null Hypothesis:** There will be no statistically significant differences in the increases in peer acceptance for pupils with low initial levels of peer acceptance in experimental and control groups.

2. Methodology

This chapter describes the methodology. It considers different research philosophies, before discussing the methodological approach and measures selected, along with the rationale underpinning these decisions. A detailed account of the research design and implementation is given. Finally, issues of reliability and validity are explored, and ethical considerations described.

2.1 Philosophical Approaches to Research

A paradigm can be described as the 'set of beliefs and practices that guide a field,' (Morgan, 2007, p.49) and provides a frame of reference from which researchers can identify their own worldview (Mertens, 2015). There is an expectation that researchers locate their research in a specific and identifiable paradigm, as this influences 'how we know, our interpretation of reality and our values and methodology' (Doyle, Brady, & Byrne, 2009, p.176).

Paradigms can be viewed as being composed of three primary elements (Healy & Perry, 2000), described by Scotland (2012) as:

- Ontology: the lens through which the world is viewed, that is, the nature of reality and what constitutes reality.

- Epistemology: the nature of knowledge (Allison & Pomeroy, 2000), that is, how we know what we know, and how knowledge is created, acquired, and communicated.

- Methodology: the strategy used in investigating the reality, and the why, what, where, when and how of data collection and analysis.

As paradigms are built upon different ontological and epistemological assumptions, with different assumptions of reality and knowledge, the paradigm

selected influences the questions which are posed by researchers and guides the methodological decisions taken in answering these questions (Scotland, 2012).

Four major paradigms applicable to research within psychology and educational psychology are described in Table 2.1.

Table 2.1 The ontology, epistemology and methodology of four research paradigms

Paradigm (basic belief)	Ontology (nature of reality)	Epistemology (nature of knowledge)	Methodology (approach to inquiry)
Positivism	Realism: a single, verifiable, external, and independent reality	Objectivist: findings are true, facts and laws can be established	Experimental. Verification of hypotheses. Quantitative
Post-positivism	Critical realism: reality exists but can only be known imperfectly and probabilistically	Modified objectivist: 'facts' and truths may be established but are only probable	Quasi-experimental. Primarily quantitative, but may include qualitative
Constructivism	Relativist: there are multiple, individually, and socially constructed realities, no universality	Subjectivist: knowledge is personal, findings are created between the researcher and participants	Primarily qualitative. Rich descriptions of contextual factors
Pragmatism	Singular and multiple realities	Knowledge is relative and provisional, and exists in different layers	Flexible methods matched to the research question and purpose. May use mixed methods

2.2 Paradigm of the Current Research

A post-positivist standpoint was adopted for the current research study.

Unlike positivism which views knowledge as having an objective and single reality (Healy & Perry, 2000), a post-positivist approach more readily allows for research in uncontrolled, real world settings (Mertens, 2015). Post-positivism acknowledges that reality cannot be perfectly detected, only within certain parameters of probability (Scotland, 2012), with findings of research always imperfect and fallible (Robson, 2011).

A post-positivist approach supported the use of quasi-experimental methodology in addressing the research question, measuring the objective impact of CL, and seeking findings which are of relevance to the real world context of both schools and educational psychology practice (Robson, 2011).

A constructivist approach viewing knowledge as subjective, personally experienced and interpreted may be of relevance in future work investigating CL in order to illuminate the lived reality of the intervention for pupils and staff through detailed descriptions of their experiences (Cohen, Manion, & Morrison, 2011). This is discussed further in section 4.4.2.

2.3 Research Methodology

2.3.1 Research Design Selected

In line with the philosophical stance and post-positivist paradigm adopted, the study employed a quasi-experimental pre-test post-test non-equivalent groups design. This was carried out across four schools, with an experimental and wait list control group class in each school.

2.3.1.1 Quasi-Experimental Designs

A quasi-experimental design is one which takes an experimental approach but where random assignment to condition is not present (Robson, 2011).

Quasi-experimental designs can be viewed as particularly useful when randomisation is impractical but where an experimental design is desirable, for example, when aiming to demonstrate causality between an educational intervention and outcome (Muijs, 2010). For the current research, true randomisation was impractical because practicalities dictated intact class groupings, and although the control group was selected to be similar to the intervention group, it was neither matched nor randomised.

In a hierarchy of quasi-experimental designs, those towards the top employ both control and experimental conditions, with pre and post-testing in both conditions (Harris et al., 2006). For this reason, designs which involved only an experimental condition or which used post-test measures only were rejected for use in the current study.

2.3.1.2 The Use of a Comparison Group

The use of a wait list comparison group was considered in addition to the wait list control group. It was envisaged that a comparison group would involve teachers increasing the amount of group work in their classrooms, but without using CL techniques. This would allow a greater insight in to whether any changes in the intervention group were likely to be due to CL techniques themselves, or to do with an increased amount of time and focus on group work.

However, the use of a wait list comparison group was rejected during the pilot and stakeholder engagement stage, as it became clear from discussions with class teachers that simply increasing the amount of time spent in group work

would be difficult in practice. Reducing the amount of group work was dismissed as not ethically sound.

2.3.1.3 Limitations of Quasi-Experimental Designs

The lack of randomisation and potential non-equivalence between groups in a quasi-experimental design poses certain difficulties. In comparison to randomised controlled trials, the use of a quasi-experimental design is often associated with an increased risk of threats to reliability and validity, and concerns that any conclusions drawn may be less certain regarding causation (Robson, 2011). Indeed, rather than being able to firmly draw conclusions, Shadish, Cook, & Campbell (2001) argue that quasi-experiments require researchers to identify a causal claim, and then examine plausible alternative explanations which may falsify the finding.

Please refer to section 2.11 for further discussion of issues of validity and reliability.

2.3.2 Alternative Designs Considered

2.3.2.1 Randomised Controlled Trials (RCTs)

RCTs are often considered the 'gold standard' experimental design with the highest level of credibility regarding causality because the process of randomisation is purported to evenly distribute both known and unknown confounding variables between the intervention and control groups (Harris et al., 2006). However, in real world research it is often not feasible or ethical to conduct RCTs (Harris et al., 2006), and whilst there may be some relative advantages to an RCT, Shadish, Cook, & Campbell (2001) argue that an RCT should not automatically be the first choice of design, even when feasible. They note that although RCTs may theoretically be stronger than other experimental designs, they are not necessarily so in practice. Indeed, with relatively small

samples, the ability of randomisation to evenly distribute confounding variables is questionable (Ju, 2009).

2.3.2.2 Single Case Experimental Design (SCED)

SCEDs have been described as a diverse and powerful set of experimental techniques (Nock, Michel, & Photos, 2007). SCEDs allow targeted and systematic analysis of the effectiveness of an intervention at an individual level through repeated measurement of a DV over time (Robson, 2011).

However, whilst a SCED design could potentially look at the effects of CL on a small number of specific children, such a design is less able to answer questions around the wider impact of an intervention. Given the gaps highlighted in the evidence base, and the resulting rationale for the research questions, it was judged that employing a SCED at this stage would be taking too narrow a view.

2.3.2.3 Non-Experimental Design

As described earlier in the chapter, a non-experimental, constructivist, qualitative approach to exploring the impact of CL may be of interest in the future. However, with a paucity of peer reviewed and published literature looking at the impact of CL on peer acceptance and perceptions of peer support in a UK mainstream population, it was judged important at this stage to build the evidence base as to whether or not CL does have a measurable impact in this context, before potentially delving deeper into an exploration of the lived experience.

2.4 Stakeholder Involvement

A number of stakeholders were considered during the planning and implementation of the current research study. The key stakeholders involved were:

• The University of Nottingham
The research was conducted as part of a requirement for the researcher's doctoral training programme (DAppEdPsy). The research focus and design was refined and agreed through discussions with tutors.

• The Local Authority
The researcher was on professional placement in the local authority where the research took place. The focus and design was agreed with the Principal EP and a Senior EP in the EPS.

• School, staff and pupils
Schools, their staff, and their pupils were central to the research. School staff received training in CL, and all parties received feedback on the outcomes of the research.

2.4.1 Stakeholder Engagement and Selection of Schools

Through discussion with the Principal EP and a Senior EP in the local authority, four schools with a degree of commonality were identified as potentially suitable for the research design. School link EPs made initial contact with SENCOs and head teachers, and all schools expressed interest in the research.

The researcher then spoke to head teachers to share further information, answer any questions, and clarify what participation would mean for the school and staff. An email was sent summarising key points (Appendix G), and arrangements made for the researcher to meet with relevant class teachers.

During the meeting with class teachers, the research study was explained, and implications of involvement discussed. An information sheet was provided (Appendix H), and written consent gained (Appendix I).

Once classes were selected, parental consent was sought. Letters to parents were sent out via the usual school procedures. The letter explained the research and asked for consent for their child to complete the measures (Appendix J).

Please refer to section 2.10 for further information on consent procedures.

2.4.2 School Characteristics

Four mainstream city primary and junior schools in an East Midlands local authority where the researcher was on placement were recruited to take part in the research.

Information from OFSTED described all the schools as larger than average. All schools had above average numbers of pupils on the SEN register, and above average numbers of pupils entitled to pupil premium funding. All schools had average, or, more commonly, above average numbers of pupils from ethnic minority backgrounds and who spoke English as an additional language. Three schools were judged by OFSTED as 'requires improvement', and one was judged as 'good' at the start of the intervention period.

School specific information is contained in Table 2.2.

Table 2.2 Key characteristics of the four participating schools

School	Size	Category and age range	Number of pupils on the SEN register	Number of pupils entitled to pupil premium	Proportion of pupils from ethnic minority backgrounds	Proportion of pupils who speak English as an additional language	OFSTED grading and date
School A	Much larger than average	Primary 3 – 11	Above average	Well above average	Above average	Above average	Requires improvement)
School B	Larger than average	Junior 7 – 11	Above average	Above average	Above average	Above average	Requires improvement
School C	Much larger than average	Primary 3 - 11	Above average	Well above average	Above average	Above average	Requires improvement
School D	Larger than average	Primary 3 - 11	Well above average	Above average	Average	Average	Good

2.5 Participants

2.5.1 Sampling Strategy

Participants were obtained through a convenience**,** cluster sample of pupils in Year 4 or 5 classrooms within schools which expressed an interest in the research.

In each school, head teachers and school staff were responsible for deciding whether Year 4 or Year 5 classes would take part in the intervention. Research specific factors such as classes where a teacher would be returning or leaving during the intervention period, job shares, and interest in taking part were taken in to consideration.

Inclusion criteria were as follows:

- Year 4 or 5, with a parallel class in the same year group
- Main class groupings not based on ability or chronological age
- One main teacher (minimum of 4 full days)
- Not anticipated that there will be any staffing changes during the intervention period
- Intact class predominantly taught by class teacher (although may be taught in sets for some lessons, e.g. maths)
- Both classes planning to cover the same curriculum content during the Summer term
- Teachers available for a ½ day training course in the Spring term
- Teachers willing and able to deliver the intervention in the Summer term
- Teachers have never had training in, or implemented, CL

Once the year group and classes were selected, class allocation to intervention or control was through a computerised coin toss. Teachers remained with their usual class, with intact, existing class groupings.

2.5.2 Class Teacher Characteristics

Table 2.3 Class teacher characteristics (intervention and control)

School	Condition	Teacher gender	Number of years teaching
School A	Intervention	Female	1 (NQT)
	Control	Female	4
School B	Intervention	Female	20
	Control	Male	3
School C	Intervention	Female	1 (NQT)
	Control	Male	8
School D	Intervention	Male	4
	Control	Female	2

There were three female and one male intervention class teachers. Two of the teachers were in their NQT year, having taught full time in their respective schools the previous year during their training. The other teachers had taught for four and twenty years. This represented a large range in the teaching experience of intervention class teachers.

There were two female and two male wait list control class teachers. Three teachers had taught for two to three years, and the fourth, for eight years.

2.5.3 Class Characteristics

Table 2.4 General information on each class (year group, number of pupils, gender and percentage of the class with consent)

School	Condition	Year Group	Number in class	Gender M/F	Percent of class with consent	No consent
School A	Intervention	5	27	M 13 F 14	85.2%	M 1 F 3
	Control	5	28	M 14 F 14	96.4%	M 1 F 0
School B	Intervention	4	25	M 12 F 13	96.0%	M 0 F 1
	Control	4	26	M 14 F 12	73.1%	M 5 F 2
School C	Intervention	5	20	M 8 F 12	65.0%	M 4 F 3
	Control	5	20	M 12 F 8	95.0%	M 1 F 0
School D	Intervention	4	29	M 17 F 12	79.3%	M 4 F 2
	Control	4	28	M 11 F 17	78.6%	M 3 F 3

Pupils were aged between 8 years 7 months and 10 years and 6 months at the start of the intervention period.

Parental consent for their child to complete the measures was gained through an opt in process, in accordance with ethical guidance. In school C, rates of

consent were particularly low in the intervention class, which meant that only half of the boys had consent. This pattern was similar in school B's control class. Teachers did not report any group similarities or differences between those who did and did not complete the measures.

The total number of pupil participants across the four schools are shown below.

Table 2.5 Total number of pupil participants in intervention and control conditions, including information on consent levels

	Intervention	**Control**	**Total *N***
***n* with consent**	83 (41m, 42f)	87 (41m, 46f)	170
***n* with no consent**	18 (9m, 9f)	15 (10m, 5f)	33
Total *N*	101 (50m, 51f)	102 (51m, 51f)	203

In total, there were 101 pupils who received CL, and 102 pupils who were part of the wait list control group. From this, 84% of pupils (N=170) had consent to complete the measures, although the individual picture between classes and schools varied.

2.6 Procedure

2.6.1 Training

Intervention class teachers received one half day researcher led training session in CL. As literature indicates that more direct approaches of CL are easier for teachers to pick up, require less training time, and are more easily implemented, the Kagan model was selected (Johnson & Johnson, 2002).

The training session included:

- an introduction to CL (including how it is distinct to group work, the main principles of CL, CL research literature regarding social, emotional, and academic outcomes, and primary psychological theories underpinning CL)
- the structured teaching of social skills using T charts, modelling, practice, reflection, and pupil self-monitoring (as well as the importance of, and examples of, small group social skills)
- group processing and the importance of reflection
- training on five main Kagan structures:
    - *timed think pair share (square)*
    - *numbered heads together / travelling heads together*
    - *quiz quiz trade*
    - *inside outside circle*
    - *robin structures (including rally robin, round robin, and simultaneous round table)*
- selecting appropriate tasks and the role of adults in the classroom
- structuring CL groups (e.g. a maximum of four pupils, with teacher selected, heterogeneous groupings which take account of academic ability, gender, social ability, and friendships)

Additionally, there was time dedicated to clarifying the research and expectations, asking questions, and an opportunity for staff to apply the Kagan structures to their own lesson plans for the forthcoming term.

Teachers were given copies of training materials, including a resource booklet 'teaching social skills and the cooperative learning structures' to support them when later planning lessons independently, and to support them in implementing the key CL structures.

Extracts from the CL training materials can be seen in Appendix K.

2.6.2 Cooperative Learning Intervention

Table 2.6 An overview timetable of the study

Pre Test	Easter holiday (2 weeks)	CL		Half term (1 week)	CL	**Post Test**
		1 week Social Skills	**6 weeks**		**4 weeks**	

The CL active period ran for 11 weeks, including an initial week of social skills training. The intervention was split by a one week half term holiday.

During the intervention period, it was expected that class teachers implemented CL approaches in at least four lessons each week. All CL sessions were delivered by usual class teachers to intact class groups on a whole class basis. The importance of not using CL strategies when teaching mixed class groups, for example, when in maths sets, was highlighted during the training session.

In school C, children were taught in their class groups by their class teacher for the whole day. In schools A, B and D, children moved in to different class groups at least once a week, and in the case of school A, daily. Although such an arrangement may have increased the risk of diffusion of treatment, it was not changed because the arrangement had been to do so all year, and to change a setup which had been in place over the course of the year may have itself impacted on post-test scores (for example, because pupils would spend more time with their class group). Furthermore, schools had organised themselves in this manner in order to provide the most suitable education for each pupil, meaning that it was deemed unethical to alter these systems.

There was an unanticipated staffing change in school B during the intervention period, where the class teacher reduced her days, teaching four days rather than five, with a new teacher taking the class for the fifth day.

2.6.3 Wait List Control

During the intervention period, wait list control classes received teaching as usual, in their usual groups, by their usual class teachers. That is, pupils in the wait list control classes received no intervention during this period.

Following the completion of the research, staff teaching in the wait list control classes were offered the opportunity to receive the same half day training course as the intervention teachers.

2.6.4 Intervention Fidelity

Intervention fidelity describes the extent to which key elements of an intervention are delivered as intended (Gearing et al., 2011). The reporting of intervention fidelity is often considered crucial in allowing others to understand the quality of a study, and in understanding factors which may have influenced the findings of a study, particularly important when evaluating the impact of an intervention (Murphy & Gutman, 2012).

Intervention fidelity can be understood through examining the following key elements:

- Adherence
- Exposure

2.6.4.1 Intervention Adherence

Adherence has been described by Carroll et al (2007) as intervention fidelity's bottom line.

Three intervention adherence checks were carried out in each intervention class. An understanding of the intervention adherence was particularly important for the current study as the intervention was implemented in four different schools, by four different teachers, thus posing a potential threat to validity, and a factor which would be important to understand when interpreting findings.

Intervention adherence was assessed using a checklist which included key facets of CL (Appendix KL).

Judgements were recorded next to each criteria as to whether they were: *present, beginnings of / some evidence, or absent*, and were scored as 2, 1, 0 respectively. Criteria which were deemed to be core elements of CL were given a double weighting, that is, positive interdependence, equal participation, small

group skills, and reflection on small group skills. There was a maximum score of 50.

Scores for each observation (1-3) in each intervention class (A-D) are shown in Table 2.7.

Table 2.7 Intervention adherence: Raw scores and percentages for each intervention class

Intervention class in	Intervention adherence check 1		check 2		check 3		Mean raw score	Mean %
	Raw	%	Raw	%	Raw	%		
School A	36	72	37	74	43	86	38.7	77%
School B	42	84	44	88	27	54	37.7	75.3%
School C	28	56	44	88	45	90	39.0	78%
School D	31	62	35	70	27	54	31.0	62%

Adherence was good overall in schools A, B, and C. Adherence was lower in school D. These aspects are discussed in more detail in section 2.6.6.

One intervention adherence check was carried out in each control class using the same checklist as used in the intervention classes. Key features of CL were not observed during paired or group work in the control classes.

In order to assess the reliability of the researcher's judgements, the third observation in each intervention class was also rated by a colleague and fellow TEP. This indicated that agreement was high, ranging from 91% to 100% agreement.

Table 2.8 Inter rater reliability scores for the third adherence check

Intervention class in	Intervention adherence check 3			% inter rater agreement
	Researcher total score	Fellow TEP total score	% adherence	
School A	43	43	86%	100%
School B	27	28	54 – 56%	96%
School C	45	41	82 – 90%	91%
School D	27	29	54 – 58%	93%

2.6.4.2 Intervention Exposure

Intervention exposure was recorded by each intervention class teacher completing a daily log of CL activities (Appendix M). In line with Carroll et al (2007), this documented frequency, duration and coverage of sessions.

All pupils were present for pre and post-test data collection, and all pupils were present for more than 80% of the CL sessions, so all data was included in analyses.

Actual intervention length was less than the anticipated 11 weeks for all schools due to the impact of residential school trips, bank holidays, OFSTED visits, school closures due to polling, and transition days/weeks.

Table 2.9 summarises intervention exposure information collected from teacher logs, along with a summary of overall adherence.

Intervention exposure and adherence for each intervention class are discussed in further detail in section 2.6.6.

Table 2.9 Intervention fidelity in the four intervention classes (adherence and exposure)

Intervention class in:	Frequency	Days with CL	Total number of CL sessions	Duration of structures	Intervention duration	Mean % adherence	Structures used (coverage)
School A	Usually twice daily	22 days	39	5 – 20 minutes	8 weeks	77%	Timed think pair share Think pair share and Squared Numbered heads together Travelling heads together Quiz Quiz trade
School B	Once or twice a week usually	18 days	22	15 – 30 minutes	10 weeks	75.3%	Think pair share Numbered heads together Travelling heads together Quiz Quiz Trade
School C	At least twice daily	23 days	47	2 – 5 minutes Usually repeated within a lesson	7 weeks	78%	Think pair share Think pair share and Squared Quiz Quiz trade Rally robin Numbered heads together
School D	Usually daily, sometimes twice daily	42 days	46	4 minutes – 1 hour (mean 12.5 minutes)	9 weeks	62%	Think pair share square Robin structures Rally robin - pairs and groups Numbered heads together Inside Outside Circle

2.6.5 Diffusion of Treatment

Class teachers completed a 'cooperative learning audit' questionnaire (Appendix N) which asked them to indicate numerically (1-7) how often they used twelve different CL structures in their teaching approaches and delivery. This was completed prior to, and following, the experimental period. The questionnaire allowed an understanding of the way in which paired and group work activities may have been structured over the research period, and for an assessment of any changes in practice (for example, in the case of treatment diffusion).

It is important to note that items were reverse scored, meaning that the greater a teacher's use of CL, the lower the score, and the lower a teacher's use of CL, the greater the score. Therefore, the maximum score of 84 would indicate that a teacher never used any of the structures. The minimum score of 12 would indicate that a teacher used each of the structures daily.

Pre and post intervention scores are shown in Table 2.10 below:

Table 2.10 Intervention and control teacher scores on the 'Cooperative Learning audit' (pre and post test)

Condition	School A		School B		School C		School D	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Intervention	72	37	75	49	72	29	63	32
Control	72	72	75	75	72	70	62	70

Data from the CL audit indicated that, at pre-test, class teachers in both the intervention and control classes were similar in their use of CL structures. The use of structures was low for all teachers. The large majority of the structures were rated as a '7' by all teachers, indicating that they never used them. The most frequently used structures were 'think pair share' and 'timed think pair

share'. Within a school, responses made by class teachers in the intervention and wait list control group appeared similar.

At post test, the use of structures for teachers in the control classes appeared broadly unchanged. That is, their scores indicating their use of CL structures before the experimental period were similar to their scores after the experimental period, suggesting little change in teacher or classroom practice from pre to post test.

However, the use of structures for teachers in the intervention classes appeared to change markedly from pre to post test, as could be expected. The greatest self-reported change was in school C, with teachers in both schools C and D reporting using CL structures with the greatest frequency.

2.6.6 A summary of each intervention class context and intervention delivery

This section summarises key information regarding the class context and delivery adherence and exposure for each intervention class.

School A

The intervention teacher in school A was female and in her NQT year. Children in this class moved to a different class grouping once a day.

Intervention adherence in school A was good. Use was very frequent on available days, often twice daily on the days when CL was used. However, there were more interruptions to available days than in schools C and D, meaning that CL was used on 39 occasions over 22 days in an eight week intervention period. Exposure was high compared to school B, but lower than that of school C and D.

When CL structures were used, they were often used in a short burst fashion several times during a lesson. Pupils appeared engaged in CL, and very quickly learned the names, processes and requirements of each structure. Pupils commented positively about CL during observations.

School B

The intervention teacher in school B was female and in her 20th year of teaching. Children in this class moved to a different class grouping at least once a week.

There was an unanticipated staffing change in school B during the intervention period, where the class teacher reduced her days, teaching four days rather than five, with a new teacher taking the class for the fifth day (in the middle of the week).

Intervention adherence in school B was good. Most elements of CL were usually present, although on occasions, there was insufficient emphasis given pre and post task to small group skills and group processing. Explicit emphasis on positive interdependence at the outset of a task was usually, but not always, present.

Although intervention adherence in school B was good, the exposure was lower than in the other three schools, with CL appearing to be the exception rather than the norm. Sessions took place only once or twice each week. In the last seven weeks of the ten week active intervention period in school B, CL was only used on nine occasions. In total, CL was only used on 22 occasions, which was less than half of that in schools C and D (47 and 46 occasions respectively). The majority of CL sessions were in the first three weeks of the intervention period, prior to the staffing change. School B had the longest intervention period (ten weeks) but use of CL was 'diluted' and infrequent over this period.

When CL was used, there was a tendency for the session to be of longer duration than in the other schools. Observations indicated that this was not

necessarily due to more CL being in place, but instead reflected a high amount of teacher talk in introducing, setting up and managing each activity. The amount of time in which pupils were able to interact freely, peer to peer, appeared lower than in the other three intervention classes. The class teacher appeared to retain a greater degree of control over interactions in school B, and, at times, class wide comments by the class teacher about certain, specific pupils during their cooperative efforts appeared negative in nature.

School C

The intervention teacher in school C was female and in her NQT year. Children were taught in their class groups by their class teacher for the whole day

Intervention adherence in school C was good, continued to improve throughout the study, and exposure was high. There was a strong emphasis on social skills and group processing. CL was used at least twice daily, characterised by short bursts, and often used multiple times within a lesson. CL became the norm in this class over a seven week intervention period, with a high number of available days and minimal timetabling disruptions. CL was used on 47 occasions over a seven week period. This was the shortest intervention period of the four schools, but use was consistently frequent.

During the first week of CL, the class teacher in school C highlighted concerns over her ability to manage to implement CL in her classroom as her class was historically quite divided, with groups of pupils who wanted to work together, and who were very vocal in their resistance to working with other pupils. She had found that this had continued across the first week of CL. However, she persevered, initially carefully grouping pupils to take into account pupils who *could* work together, and praising them in their use of specific social skills and small group skills. Over the next few weeks, information from the class teacher and from researcher observations indicated that pupils became used to a pattern of working cooperatively, working with a range of other pupils, and did so both effectively and without any apparent resistance. Observations during adherence checks two and three suggested that a clearly observable positive

and celebratory atmosphere had developed in the classroom as the intervention period progressed, with pupils clapping, ‘high 5-ing’ and praising their partners and group. This was quite a marked change in atmosphere from prior to the intervention, and the first adherence check.

School D

The intervention teacher in school D was male and in his fourth year of teaching. Children in this class moved to a different class grouping at least twice a week.

Intervention adherence in school D was variable during adherence observations. In this class, key elements of CL were usually present and adherence good, but, at times, some elements were poorly introduced. As such, although both the class teacher and the pupils appeared enthusiastic, the intervention adherence percentage was lowest in this class. However, there was always a strong focus on collaboration and child led discussion, rather than teacher led learning, as well as social skills and group processing. Pupil-pupil interaction and dialogue was high, although equal participation was not necessarily always present.

CL was used daily, and sometimes twice daily over a nine week intervention period. There were minimal disruptions to this pattern and a high number of available days, with CL used on 46 occasions over 42 days. Although reported CL sessions ranged from four minutes to one hour in duration, the mean duration was 12.5 minutes.

2.7 Measures

In order to answer the research questions, two dependent variable measures were used:

- the Classroom Supportiveness Scale (CSS) was used to measure perceptions of peer support (Battistich, Solomon, Kim et al., 1995).

- the Social Inclusion Survey (SIS) was used to measure peer acceptance (Frederickson & Furnham, 1998).

Both measures were pupil self-report. This was a straight forward, inexpensive and time efficient way of collecting data across four schools and eight different classes.

The primary focus of the current research was to assess change from the perspective of pupils themselves. As such, measures were not administered to parents or school staff.

Although triangulation of data is often considered useful, the utility and validity of teacher reports at this point was questionable as literature indicates that when asked to estimate peer acceptance of pupils in their class, teachers may be unduly influenced by knowledge of the pupils' academic levels and gender (Parker & Asher, 1987). Thus, pupil responses are generally considered more diagnostically efficient (McKown, Gumbiner, & Johnson, 2011).

Nevertheless, difficulties with self-report data should be acknowledged, including participant error and participant bias whereby factors such as emotional affect, motivation or social desirability bias may influence pupils' responses on the measures.

2.7.1 Measuring Perceptions of Peer Support: The Classroom Supportiveness Scale (CSS)

Perceptions of peer support were measured using the CSS, taken from the primary aged version of the Sense of School Community Scale (Battistich, Solomon, Kim et al., 1995).

The primary aged Sense of School Community Scale consists of 38 items presented in three subscales: Classroom Supportiveness (CSS), School Supportiveness, and Classroom Autonomy (Battistich, Solomon, Kim et al., 1995). Used as a standalone measure, the CSS consists of 18 items, which aim to assess the extent to which pupils feel their classmates to be supportive and helpful (Appendix O). Children respond to each item using a five point Likert scale ranging from 1 (disagree a lot) to 5 (agree a lot).

The two remaining subscales (School Supportiveness, and Classroom Autonomy) were not deemed relevant to either the research questions or to the areas in which CL may be expected to impact in the current study.

Literature indicates a high internal consistency for the overall Sense of School Community Scale (*a* = .91), and for the CSS (*a* = .85) (Battistich, Schaps et al., 2000). Additionally, the scales are reported to be equally reliable for boys, girls, and pupils from different ethnic backgrounds (Roberts, Horn, & Battistich, 1995).

2.7.1.1 Rationale for Using the Classroom Supportiveness Scale

The CSS (Battistich, Solomon, Kim et al., 1995) is able to assess pupils' perceptions of support by focusing on pupil rather than teacher or parent perspectives.

Furthermore, as a standalone scale, it focuses only on perceptions within the class group, where CL may have had an influence, as opposed to the wider

school settings where, in the current research design, it would not be expected that CL would have an impact.

A common concern with self-report measures is the tendency for people to agree (rather than disagree) with statements, termed response acquiescence set or response bias (Coolican, 2009). The use of reversed and negative items on the CSS may go some way to addressing this.

Consideration was given to using the Classroom Life Instrument (Johnson, Johnson, & Anderson, 1983), as used by the three perceptions of peer support studies in the systematic review e.g. Bertucci et al. (2010). However, it was judged that the benefits of potentially being able to distinguish between academic and personal support was outweighed by lower reliability indices, an absence of reversed items on these scales, and a reliance on only nine items if the scales were used in isolation. If used as a full scale, there were a high number of items which were deemed inappropriate to the current study.

2.7.1.2 Scoring the CSS

The CSS (Battistich, Solomon, Kim et al., 1995) was scored by hand, taking into account the six reverse scored items.

2.7.2 Measuring Peer Acceptance: The Social Inclusion Survey (SIS)

Peer acceptance was measured using the SIS (Frederickson & Furnham, 1998).

The SIS is a forced choice sociometric scale designed to be administered on a whole class basis to children aged 7 years and above (Frederickson & Furnham, 1998).

Each pupil was given two questionnaires (one for work, one for play) with the names of children in their class listed down the left hand side (Appendix P).

Pupils were asked to indicate next to each name how much they liked to work or play at school with each child in their class by ticking a smiling face (like to work/play), sad face (prefer not to work/play), straight mouthed face (don't mind either way) or a question mark if they did not know the child well enough to decide (Frederickson & Furnham, 1998).

The SIS calculates peer acceptance using the judgements made by peers about an individual child (see example in 2.7.2.2). Therefore, pupils were asked to cross out their own name on the questionnaires. This meant that they were unable to rate themselves, and did not count as a respondent when scores were calculated.

Content validity for the SIS is enhanced through assessing peer acceptance in both the play and work contexts, along with the use of both positive *and* negative nominations (Coie, Dodge, & Coppotelli, 1982). This avoids difficulties associated with unidimensional techniques such as rating methods (Gifford-Smith & Brownell, 2003).

Although Bukowskil & Hoza (1989) note that the stability of many sociometric measures are 'not overwhelmingly impressive' (p.28), Frederickson & Furnham (1998b) reported Pearson's product moment correlations of between .70 to .78 for test retest reliability of the 'like to play' scale over five weeks, a forced choice measure comparable to the SIS. This was one of the best scores from the methods they investigated.

2.7.2.1 Rationale for Using the SIS

The SIS (Frederickson & Furnham, 1998) was selected as a measure of peer acceptance because it is a quick and easy method of collecting sociometric information.

The use of a forced choice technique meant that information was gained about each pupil in a class, enabling broader pictures of acceptance, whereas limited nomination procedures risk only highlighting those pupils who are most prominent (Rubin, Begle, & McDonald, 2012).

2.7.2.2 Scoring the SIS

In line with the weighted scores approach adopted by Andre et al. (2011) and Craig (2010), peer acceptance scoring for each child was attained by subtracting the number of sad faces received from the number of happy faces. In order to enable a comparison to be carried out between classes of different sizes, and different numbers of male and female respondents, this was then divided by the total number of respondents, and multiplied by 100. The weighted scores approach led to interval level data.

In doing so, each member of the class received four peer acceptance scores, according to four different contexts:

- work same gender WSG
- work different gender WDG
- play same gender PSG
- play different gender PDG

An illustration of this process would be a female pupil who received seven smiling faces, three neutral faces, two sad faces, and zero question marks from her twelve female peers in the class in response to the question, ‘how much do you like to work with…’. Her peer acceptance score would be calculated as seven smiling faces minus two sad faces, giving a score of five. This would then be divided by twelve (number of respondents), and multiplied by 100 giving a final work same gender (WSG) peer acceptance score of 41.67.

If she had received twelve smiling faces from all twelve respondents, her peer acceptance score would have been calculated as twelve smiling faces minus

zero sad faces, giving a score of twelve. This would then be divided by twelve (number of respondents), and multiplied by 100 giving a final work same gender (WSG) peer acceptance score of 100.

On the other hand, if she had received twelve sad faces from all twelve respondents, her final work same gender (WSG) peer acceptance score would have been -100.

Thus, the peer acceptance scores for pupils on each context (WSG, WDG, PSG, PDG) each ranged from -100 (low peer acceptance) to +100 (high peer acceptance).

2.8 Data Collection Procedures

Pupils completed each measure at pre and post-test.

Pre-measures were taken in the week immediately prior to the Easter holiday, and the post measures were taken in the week immediately following the intervention. This meant that there was a 14 to 15 week duration between pre and post measures in each school. Please refer to table 2.6 in section 2.6.2 for further information.

2.8.1 Procedure for Administering Measures

A pupil information sheet was provided at each data collection point, and read out by the researcher (Appendix Q).

If parental consent had been received, children then completed a pupil 'willingness to participate' form indicating whether they were or were not willing to complete the measures (Appendix R).

Each measure was administered on a whole class basis by the researcher (SIS, then CSS). Whole class delivery was selected as practical given the staffing and time constraints of the study. Teaching staff remained in the classroom, and were available to support any children who encountered difficulties in completing the questionnaires.

Baldwin (1999) highlights the importance of ensuring uniformity of data collection procedures, and as such, a researcher designed script was used to ensure standardised delivery of measures across classes and schools and reduce researcher bias.

On the SIS (Frederickson & Furnham, 1998), children worked at their own speed but were told to ask an adult if there was a name which they were unable to read. For the CSS (Battistich, Solomon, Kim et al., 1995), the researcher read out each item and asked children to work at the same speed.

Concerns are often raised regarding bias in self-reports (Baldwin, 1999), as responses may be influenced intentionally or unintentionally by a tendency for people to select socially desirable responses in an attempt to show themselves in a good light or to please the researcher (Coolican, 2009). To minimise this, pupils were reminded to be honest, and that there were no right or wrong responses. Pupils were also informed that their responses would be not be shared, and, as pupils were not known to the researcher, that the researcher would not be able to match names to individuals once the measures were collected in.

Pupils completed the measures individually and were asked not to share their responses with any other pupils in the class, with it explained that it might upset children if they knew that someone had put a sad face next to their name. Seating in the classroom was arranged so that pupils could not see the responses that those sat near them made. Many children chose to use a book to increase their privacy.

When children had completed both measures, they were asked to take a couple of minutes to reflect privately on a friendship in or out of school which they had which was positive for them, and something which they liked about being in their class. This reflection was used to encourage all children to end the data collection on a positive note.

2.9 Pilot Study

A pilot study was carried out in the Spring term 2016 in order to trial the training materials, consider any practicalities of intervention implementation within a real classroom situation, refine the fidelity checklist, and familiarise with administration of the SIS (Frederickson & Furnham, 1998).

The pilot school was single form entry, and therefore not eligible to participate in the main study. Initial contact was through the school's link EP, with consent processes as described for the main study in section 2.4.1.

The Year 5 class teacher received half a day CL training, and implemented the approaches in her classroom over a four-week period. The researcher conducted two fidelity checks in the intervention class and one in the control class (Year 4), as well as administering the SIS once with both classes.

A comparison group in the Year 3 classroom was initially planned, where the amount of time spent in usual group and paired work activities would be increased. However, discussion with the three staff members prior to the start of the pilot study indicated that this would not be realistic as all felt that they were already carrying out a high amount of paired and group work, and therefore would not be able to increase this in any substantial way. Following similar discussions with class teachers in the main study, a comparison group was not used at either point.

The pilot was successful in supporting preparation for the main study. Key findings are outlined below:

- Whole class administration of the SIS was effective, with pupils completing the measure independently. The pilot highlighted a need to check that children had completed an entry for all names on the SIS. There was no evidence of distress caused by completing the measures.

- It was judged that adding a further measure for the main study would not place unreasonable demands on pupils or staff.

- Training was rated as highly effective by the intervention class teacher, with half a day judged to be an appropriate amount of time.

- The class teacher felt that the range of CL interventions provided through the training was appropriate, and supported her in flexibly matching CL structures to curriculum activities. The number of structures was kept the same for the main study. Indications were that the pupils quickly picked up the different structures.

- Intervention fidelity was very good in the intervention classroom. There were no elements of CL observed in the control classroom. The fidelity checklist was judged suitable for purpose.

- Observations in the intervention classroom raised issues around the role of adults, including teaching assistants during CL. Observations indicated that when adults were involved with a group, the dynamics changed and the dialogue decreased, from pupils taking it in turns and talking as a group, to pupils talking one after another directly to the adult.

  A clearer focus on the role of adults during CL was emphasised in the main study training session, with the teacher a ‘guide on the side’ not a ‘sage on the stage’, as proposed by Johnson & Johnson (1999).

- Returns of consent slips took longer than anticipated, and returns were low at just over 50%. This highlighted the need for an extended period to gain consent slips in the main study. It also highlighted a concern, reality and potential limitation of real word research. Discussions held with university tutors at this stage indicated that despite these potential difficulties, opt out consent would not be ethically acceptable.

2.10 Ethical Considerations

Key ethical guidelines from the BPS (BPS, 2009; BPS, 2010; DECP, 2002) and the Health and Care Professions Council (HCPC, 2012) were considered, and ethical approval obtained from the University of Nottingham's ethics committee (Appendix S).

Ethical considerations pertinent to the current research are described below, as are the steps taken to minimise such concerns:

Informed Consent: Informed consent was gained at several levels, with processes designed to ensure voluntary participation, the provision of adequate information, and opportunities for those involved to ask questions.

CL was viewed as falling within the range of usual curriculum activities and as part of school practice rather than an additional curriculum or withdrawal to a specific intervention. As such, and in line with the BPS (2010) Code of Human Research Ethics, approval and access granted by head teachers was deemed sufficient for pupils to be part of a CL intervention class.

Opt in consent was gained from parents to allow their child to complete the measures (Appendix J). Children for whom parental consent was not gained did not complete the measures.

Pupils provided written consent through a 'willingness to participate' form (Appendix R).

The Right to Withdraw: School staff, parents and pupils were made aware that participation was voluntary, and that they could withdraw from the research at any point without prejudice, without being questioned, and without having to provide a reason.

It was made clear that consent could be withdrawn retrospectively, and that any requests for data to be destroyed would be complied with.

Confidentiality: Individual responses on the measures were not shared with staff, parents or other pupils. Data has been reported carefully so as to reduce the likelihood of individual pupils, classes or schools being identified.

Protection from Potential Harm: Literature indicates that a focus on peer acceptance and relationships per se does not impact negatively on children, and that the use of measures involving positive and negative peer nominations does not cause any harm greater than children might encounter in daily life (Iverson, Barton and Iverson, 1997). As such, completion of the measures was deemed unlikely to change participants' mood, be aversive or stressful.

Deception: The purpose of the research was explained to schools, parents and pupils at the outset. At no point were any stakeholders deceived.

Debriefing: Following the completion of the research, letters were given to schools, parents and pupils thanking them for their contribution to the research and outlining the key findings of the study (Appendix T).

2.11 Validity and Reliability

Within the post-positivist paradigm, researchers need to eliminate possible alternative explanations in order to claim that one variable caused a change in another, remembering that findings are 'true' only within certain parameters of

probability, and within the conditions which were encountered during the study (Mertens, 2015).

As such, consideration of threats to reliability and validity particularly pertinent to the current research are discussed below, as are the steps taken to address and/or minimise such concerns.

2.11.1 Issues of Validity

Validity refers to the accuracy of a result, that is the degree to which what has actually been measured is the same as what is claimed to have been measured (Robson, 2011). Gronlund (1981) argues that rather than discussing validity as an absolute state, it should be seen as a matter of degrees, where it is accepted that research will never be 100% valid, and potential areas of error are acknowledged (Cohen, Manion, & Morrison, 2011). In quantitative designs, this leads towards an exploration of internal and external validity (Mertens, 2015).

Concerns most pertinent to the current research study are described below.

2.11.1.1 Internal Validity

Internal validity means that 'the changes observed in the dependent variable are due to the effect of the independent variable, not to some other unintended variables' (Mertens, 2015, p.130). Threats to internal validity noted as particularly relevant to this study are discussed below.

History: changes over time due to exposure to external events.
Discussions with stakeholders prior to the intervention period did not highlight any school wide changes or additional interventions which could potentially account for changes in the DVs. However, after the half term holiday, the intervention teacher in school B reduced her teaching days from five to four days. This was not anticipated, and may have impacted on findings.

Statistical regression: change caused by extreme scores regressing towards the mean.

Barnett, Van der Pols, & Dobson (2005) note that regression to the mean is a common statistical phenomenon, and should always be considered as a possible cause of an observed change when assessing the effectiveness of an intervention.

Selection: existing differences between groups and their individuals.

Due to non-random assignment, class groups may have been different prior to the study, and therefore any existing differences between classes (and their class teachers) may have affected the outcome of the study – which could mean concluding CL did not make a difference, when it did, or vice versa (Trochim, Donnelly, & Arora, 2016). As such, inferences about causality are less certain. Similarly, non-participation can introduce selection bias (McKown, Gumbiner, & Johnson, 2011)**.**

Diffusion of treatment: contamination of intervention strategies, techniques and gains to the wait list control group.

The risk of diffusion of treatment was present in this study as control groups were parallel classes.

A pre and post-test CL audit completed by teachers provided an insight into the ways in which paired and group work activities were structured over the research period, and allowed an assessment of any changes in practice (see section 2.6.5). Teacher responses on the audit suggested that all teachers were similar at pre-test, with low use of CL structures. At post-test, responses indicated that teaching approaches had changed in the intervention classes, with teachers using a range of structures frequently. No such change was observed in responses from the control group teachers.

Furthermore, intervention fidelity observations carried out in control classes did not identify any CL.

However, in three schools, some children spent at least some of their week with the control group class teacher, and vice versa. As such, it is possible that intervention teachers may have inadvertently used some of their developing strategies when teaching a mixed class set of intervention and control pupils. Diffusion may also have occurred between the children, with any increased skills in the intervention class children impacting on the skills of the control group children when in a mixed set.

2.11.1.2 External Validity

External validity refers to the extent to which the findings of the research can be applied to other contexts, situations, times, and populations (Cohen, Manion, & Morrison, 2011).

In the current research, ecological validity is raised through the use pre-existing classes, with their usual teachers, and the usual curriculum. Rather than selecting schools or teachers as particularly effective or ineffective, or employing an unrealistic amount of support in order to produce a positive finding, research was conducted in such a manner as to establish the impact of CL within usual, mainstream, local authority schools, with realistic levels of support and training.

As the study was carried out across four schools, this increases the external validity of its findings, and reduces the chance that any findings are associated with characteristics of a specific teacher. It could be reasonably assumed that the findings could be applied to similar populations in similar contexts.

Nevertheless, the population and context remains restricted to two form entry schools in a single authority with participants drawn from Years 4 and 5 only.

Pre-established measures which had strong indicators of validity and reliability were selected. However, it is possible that the choice of DVs and the specific

measures employed may have impacted on findings and conclusions regarding the impact of CL.

Novelty and disruption effects whereby positive results arise because an intervention is new, or negative results occur because the intervention initially causes disruption may have impacted on findings. One of the intervention class teachers expressed concern about the way the intervention unsettled the class initially, suggesting that such an effect may have been active.

2.11.1.3 Further Threats to Validity

Cohen, Manion, & Morrison (2011) highlight intervention fidelity and strength of the experimental treatment as particularly crucial to consider when evaluating educational and psychological research.

Intervention fidelity:
Intervention fidelity can be described as the extent to which an intervention was delivered as intended (Mertens, 2015). Indeed, rather than jumping to 'did it work?', there is a need to understand exactly what the reality of 'it' was in any given research context (Erikson & Gutierrez, 2002).

All intervention staff received specific and focused training in using CL approaches. However, lasting only half a day, this training was not as high in duration as some of the other research literature. This may have impacted on the delivery of the intervention.

Data from fidelity observations and teacher logs provided useful information on the nature of the CL intervention in each classroom. The use of a structured checklist and inter rater reliability checks supported the judgements made. Checks did indicate that fidelity was generally good – although not all CL elements were always present (see section 2.6.4.1).

Strength of experimental treatment:

The 11 week intervention period was a pragmatic choice, linked to term length in UK schools and study timescales. Due to further school specific demands, this was not achieved in any class (see section 2.6.4.2). It is important to note that Joyce & Showers (2002) argue that an intervention of medium complexity takes teachers 20-25 trials over eight to ten weeks to become famiilar with, and Cohen, Manion, & Morrison (2011) suggest that it may not be reasonable to expect to see changes in pupils' learning, attitudes or self-concepts over a short time scale. Conclusions regarding the impact of CL must take these issues in to consideration.

2.11.2 Reliability

Reliability is concerned with stability and consistency, that is, if the research was carried out with a similar group of pupils in a similar context, would similar results be found (Cohen, Manion, & Morrison, 2011)?

The use of a script ensured standardised delivery and administration of measures between classes and schools, and both measures were suitable for primary aged pupils.

The majority of items on the CSS included statements such as 'Pupils in my class are willing to go out of their way to help someone', whereby pupils would need to select 'agree a lot' if they wanted to indicate the most positive response. However, on reversed items such as 'Pupils in my class don't get along together very well', pupils would need to select '*disagree* a lot' if they wanted to indicate the most positive response.

It is possible that some pupils may have had difficulty in understanding the reversed items on the CSS, and difficulty in selecting their actual intended response. However, the effects of this were minimised by asking pupils to work at the same speed as the researcher, clarification of each item being given by the researcher, and having additional staff available to help children.

It is also possible that participant error and participant bias were threats in the current study. That is, pupils' responses may have fluctuated from pre to post-test due to influences other than the intervention, such as emotional affect, motivation or social desirability bias. This may have been heightened due to the age of the pupils and the focus of the DVs.

3. Results

This chapter presents the findings of the study, and outlines the statistical analyses involved in judging significance of the results.

3.1 Parametric and Non-Parametric Analyses

Conducting an inferential test allows the statistical significance of any findings to be determined, providing information on the probability of the results being found if the null hypothesis was true (Coolican, 2009). A statistically significant result at $p<0.05$ indicates that chance is an unlikely explanation (Kirk, 1999), and provides support to the experimental hypothesis, with reasonable confidence that a genuine effect has been identified (Coolican, 2009).

There are two main categories of inferential statistical analyses: parametric, and non-parametric (Cohen, Manion, & Morrison, 2011).

Dancey & Reidy (2014) write that for the use of parametric tests, data should meet the following assumptions:

- Interval or ratio level data
- Homogeneity of variances
- Normally distributed, that is, not considerably skewed or kurtotic
- No extreme outliers present

A number of checks were made on the pre-test data in order to decide which inferential analyses would be most appropriate.

Interval or ratio level data

All data collected was interval level. This assumption was satisfied in all cases.

Homogeneity of variances

In the circumstances where parametric tests were deemed appropriate, Levene's test for equality of variances was carried out, and accounted for in the reporting of significance (Brace, Kemp, & Snelgar, 2012).

Normal distribution:

The normality of distribution was assessed through a range of approaches, including:

- Visual analysis of histograms and Q-Q plots (Robson, 2011)
- Particular consideration of Shapiro-Wilk significance in samples up to N=50 (Hinton, 2014)
- Particular consideration of skewness and kurtosis in samples over N=50 (Brace, Kemp, & Snelgar, 2012). Skewness and kurtosis values between -1 and +1 were accepted as supporting a normal distribution (Dancey & Reidy, 2014)
- Assessing the discrepancy between mean and median values (Robson, 2011)

Outliers

Visual analysis of histograms and box plots was used to identify any extreme outliers.

It was expected that due to the nature of the DVs, that there would be some outliers, and that as suggested by Field (2013), these outliers would be interesting in their own right.

There are a multitude of approaches to managing outliers, ranging from trimming or transforming the data, winsorizing or bootstrapping (Field, 2013). As the outliers were important as part of the differential impact aspect of the research questions, these were kept in their original form. In line with Field (2013), non-parametric tests were used when data sets included multiple outliers or included one particularly extreme outlier.

3.2 Parametric Test Selection

In pre-test post-test designs where random allocation has not been possible, it is likely that the initial scores for experimental and control groups may differ. In these quasi-experimental designs with non equivalent groups, Wright (2006) highlights two main approaches to analysis:

- t tests or one way ANOVAs on the change scores (post-test score minus pre-test score)
- ANCOVA using the pre-test score as a covariate

However, literature describes 'the Lord's paradox' where the two methods can show opposite effects and lead to different conclusions (Wright, 2003). The resultant debate regarding analytical models reveals disagreement between commentators, with Senn (2006) arguing that change scores yield more biased estimates, whereas Allison (1990) argues that change scores yield *less* biased estimates, with assumptions that ANCOVAs are more powerful resting on untenable assumptions.

Wright (2006) notes that the two methods address slightly different research questions, with change scores better suited than ANCOVA to answering the question as to whether any change which occurs is on average greater for the experimental than control group.

Therefore, given the current research questions, when parametric test assumptions were met, t tests or one way ANOVAs on change scores were used. When statistically significant differences were found during ANOVAs, Games-Howell was used post hoc to identify where the differences lay. Games-Howell was selected as it provides good performance when sample sizes are unequal (Field, 2013), particularly relevant when investigating the differential impact.

In line with Wright (2006), the use of ANCOVA was deemed less suitable in addressing the current research question. Indeed, Miller & Chapman (2001)

argue that despite its ongoing use, literature roundly condemns the use of ANCOVA as a way of circumventing the interpretative problems which may arise when two groups differ at pre-treatment in a quasi-experimental design. When treatment groups differ on the covariate, for example, initial peer acceptance, Miller & Chapman (2001) argue that putting the covariate into ANCOVA will not control for or balance those differences, and when the covariate and experimental effect are not independent, interpretation of ANCOVA is seriously compromised (Wildt & Ahtola, 1978).

As a final note, some commentators suggest mixed factorial or repeated measures ANOVAs as an approach to analysis in pre-test post-test designs (Knapp & Schafer, 2009). However, the 'F' value for the treatment main effect can be very conservative (Dimitrov & Rumrill, 2003), and lead to potentially significant results being overlooked.

3.3 Non Parametric Test Selection

Non parametric tests make fewer assumptions about a data set (Field, 2013). Whilst they do still have distributional assumptions, they do not assume normality. The use of ranks manages the presence of outliers and skew (Field, 2013).

When parametric test assumptions were not met, non-parametric statistical analyses were conducted using the Independent Samples Mann-Whitney U, Wilcoxon Signed Rank, and Kruskal-Wallis tests.

3.4 Grouping Data from Multiple Schools

As data from the four schools were to be combined for analyses, it was important to establish the validity of treating them as one overall data set.

A one-way ANOVA indicated that the data from the four schools were not statistically significantly different at pre-test on the:

CSS	(F (3, 166) = 2.61, p = 0.053)
SIS work same gender	(F (3, 199) = 1.35, p = 0.260)
SIS work different gender	(F (3, 199) = 2.50, p = 0.061)
SIS play same gender	(F (3, 199) = 0.78, p = 0.508)

The data from the four schools was statistically significantly different at pre-test on:

SIS play different gender (F (3, 199) = 7.32, p = 0.000)

Inspection of the means plots and Games-Howell indicated that this difference was associated with School C. Excluding school C from the analysis on this variable was considered. However, on balance, it was judged that combining all four schools would best recognise the real world variation between schools and classes who may use CL, with interpretations taking account of this grouping.

3.5 Investigating the Differential Impact

In order to address research questions two and four, pupils' data was divided in to three groups at the analysis stage, according to their initial levels of peer acceptance and perceptions of peer support. Groups were calculated using the interquartile range:

low	below 25th percentile
medium	at or above 25th to at or below 75th percentile
high	above 75th percentile

Two approaches were taken in understanding the differential impact of Cooperative Learning on pupils' peer acceptance and perceptions of peer support.

Firstly, analyses compared changes in scores for pupils in the intervention condition with initially low, medium, and high levels of perceived peer support

and peer acceptance. This was important in understanding whether any impacts of Cooperative Learning would act so as to close the gap for potentially vulnerable pupils with low perceptions of peer support and low peer acceptance – or whether impacts would lead to greater gains for pupils with initially medium and high perceptions of peer support and peer acceptance, resulting in concerns around a widening gap which may disadvantage the most vulnerable group.

However, it was acknowledged that due to their positions at either end of the distribution, comparing high and low groups could attenuate the risks associated with regression to the mean, whereby it is more likely for low scorers to increase, and high scorers to decrease from pre to post test, regardless of any impact of an intervention. Increases from pre to post test for low pupils could be inflated, and changes for high pupils depressed due to potential ceiling effects. Thus, it was acknowledged that any increases for low groups in comparison to high groups would need to be interpreted with some caution.

Schagen & Hodgen (2009) argue that when trying to understand impact of an intervention for pupils at either end of the distribution, that is, in either low or high groups, it is essential to compare them with equivalent pupils not receiving the intervention, rather than solely with other pupils in the class.

Therefore, further analyses compared changes in scores for pupils with initially low levels of perceived peer support and peer acceptance in the intervention and control conditions.

3.6 Effect Sizes

Commentators warn of an over reliance on statistical significance (Cohen, Manion, & Morrison, 2011), and a wariness towards the 'all-or-nothing thinking' which may accompany hypothesis testing and *p* values (Field, 2013, p.85). Indeed, Schmidt & Hunter (2002) argue that 'significance tests are a disastrous method for testing hypotheses' (p.65) as they can be biased towards retaining

the null (Harcum, 1990), with non significance, at times, an artefact of low power (Field & Hole, 2003).

Instead, Lakens (2013) writes that effect sizes are the most important outcome of empirical studies, as they can help in understanding and interpreting under or over powered studies (Grissom & Kim, 2011). Effect size can be described as 'an objective and standardized measure of the magnitude of the observed effect' (Field, 2013, p.79) and allows a distinction to be drawn between statistical and practical significance (Dancey & Reidy, 2014).

Indeed, Higgins, Kokotsaki, & Coe (2012) note that 'a 'small' effect may be educationally important if, for example, it is easy or cheap to attain or is achievable with groups who are otherwise hard to influence' (p.8). Similarly, they note that even large effect sizes may be relatively less important if they have arisen from research which has optimised the effect of an intervention, rather than one which is realistic for class teachers to replicate given typical school and educational circumstances.

Following the synthesis of more than 800 meta-analyses, Hattie (2009) argued that only effect sizes greater than $d = .4$ are in the 'zone of desired effects', and indicate educational interventions which appear to be having an impact on pupil outcomes beyond typical teacher or developmental effects.

In the current study, 'Cohen's *d*' is reported following t tests, 'Eta Squared' following ANOVAs, and 'Pearson's *r*' following non-parametric tests. Where *r* is reported, it indicates the effect size for the intervention group change pre to post-test, according to the Wilcoxon Signed Rank test. Tables 3.1, 3.2, and 3.3 summarise the effect size descriptors used in the study.

Table 3.1 Cohen's *d* levels and effect size descriptors

Cohen's *d*	Effect size descriptor taken from Cohen (1988) and Field (2013)
.2	Small
.5	Medium
.8	Large

Table 3.2 Pearson's *r* levels and effect size descriptors

Pearson's *r*	Effect size descriptor taken from Field (2013)
.10	Small
.30	Medium
.50	Large

Table 3.3 Eta Squared levels and effect size descriptors

Eta Squared	Effect size descriptor taken from Field (2013)
.01	Small
.06	Moderate
.14	Large

3.7 Power

The power of a test is its 'ability to detect a significant effect, where one exists' (Dancey & Reidy, 2014, p.244). A power value of .8 when alpha is set at .05 gives confidence that there is sufficient power to detect any effects (Field, 2013). If power falls below .8, there may be a need to replicate the research with more participants in order to increase the power.

A priori power analyses using G Power software (Buchner, Erdfelder, Faul, & Lang, 2016) indicated a need for a total sample of 128 participants in order to achieve an 80% chance of identifying medium effects (d = .5) when alpha was .05. This level of participants was met for all analyses when the schools were combined, meaning that the study had sufficient power to detect an effect if one existed.

However, when analysing individual school data, power was low. Post hoc analyses indicated that power to identify medium effects for individual schools ranged from .27 to .46. This does not reach the level identified by Field (2013) and suggests up to a 73% chance of failing to detect a genuine effect. The chance of failing to detect a genuine but small effect, would be even greater.

Therefore, primary analyses were carried out on data sets as a whole, with all schools combined. Individual school analyses were supplementary, in order to provide a more detailed picture of any school specific trends, particularly important given the differences in intervention implementation and exposure between the four schools.

With regards to exploring the differential impact within the all schools combined intervention condition, power was not sufficient to identify a medium effect. Post hoc analyses indicated that the power of a one way ANOVA to identify a medium effect at p = .05, was only .50 (CSS intervention group) and .59 (SIS intervention groups).

Power was also low when exploring the differential impact between low group pupils in intervention and control conditions. For an illustration, with 25 intervention group pupils and 25 control group pupils in a sample, the power of an independent samples t test to detect a medium effect (d = .5) at p = .05 was only .41. Indeed, the power to detect a large effect (d = .8) was still only .79.

3.8 Bonferroni's Correction

The Bonferroni correction can be applied in order to control for the increased possibility of Type 1 errors occurring when multiple significance tests are used (Field, 2013). However, when lots of tests are performed, the correction is often criticised as being too strict, as by applying Bonferroni's correction, the probability of Type II errors increases as the tests have less power to detect effects (Field, 2013).

A balanced view between those who argue that Bonferroni should never be used (e.g. Rothman, 1990), and those who argue that it should always be used (e.g. Ottenbacher, 1998), is the literature arguing that Bonferroni may be applicable in some circumstances, but should never be applied in a routine or uncritical manner (Armstrong, 2014). It appears most appropriate when there are concerns about 'fishing' for significance, that is, when a large number of tests are carried out without pre-planned hypotheses (Armstrong, 2014). However, when hypotheses have been specified beforehand, then Bonferroni should be avoided as it becomes an over correction (Streiner & Norman, 2011). In this latter case, where multiple t tests have been used and the results of individual tests are important, the exact p values for each individual test should be used (Armstrong, 2014).

In light of this literature, and the increased risk of Type II errors, the Bonferroni correction was not applied in the current study. Furthermore, given the low statistical power described in section 3.7, the chance of obtaining a Type I error was judged to be relatively low. However, the potential for multiplicity was

recognised when interpreting and discussing statistically significant findings, as advised by Streiner & Norman (2011).

3.9 Perceptions of Peer Support

Perceptions of peer support were measured using the Classroom Supportiveness Scale (Battistich, Solomon, Kim et al., 1995).

Research Question 1: What is the impact of CL on pupils' perceptions of peer support?

Experimental Hypothesis: There will be a statistically significantly greater increase in pupils' perceptions of peer support following CL (experimental condition), when compared to pupils in a control condition.

In the main analysis, data from all four schools was combined. Data was normally distributed. Mean and change score data is provided in Table 3.4.

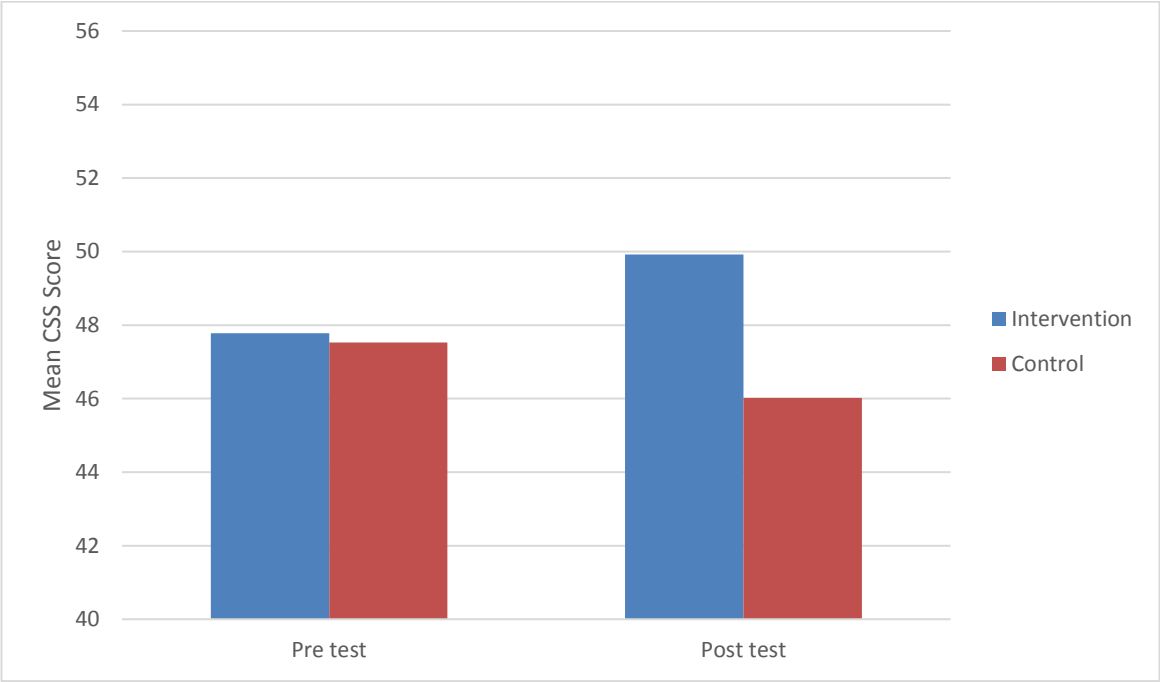
Table 3.4. Perceptions of peer support: Pre and post-test overall means and change scores for intervention and control conditions (CSS)

		N	Mean	*SD*	Change Score	*SD*
Intervention	Pre Test	83	47.78	9.64		
					2.23	8.48
	Post Test	83	49.92	10.40		
Control	Pre Test	87	47.53	9.14		
					-1.49	9.12
	Post Test	87	46.02	9.96		

Intervention and control group pre-test mean scores were very similar. An independent samples t test indicated that they were not statistically significantly different (t (168) = 0.18, p = .860).

Mean scores increased for the intervention group from pre-test to post-test, and decreased for the control group over the same period. This is illustrated in Figure 3.1.

Figure 3.1 A graph displaying overall mean CSS scores at pre and post-test for intervention and control conditions



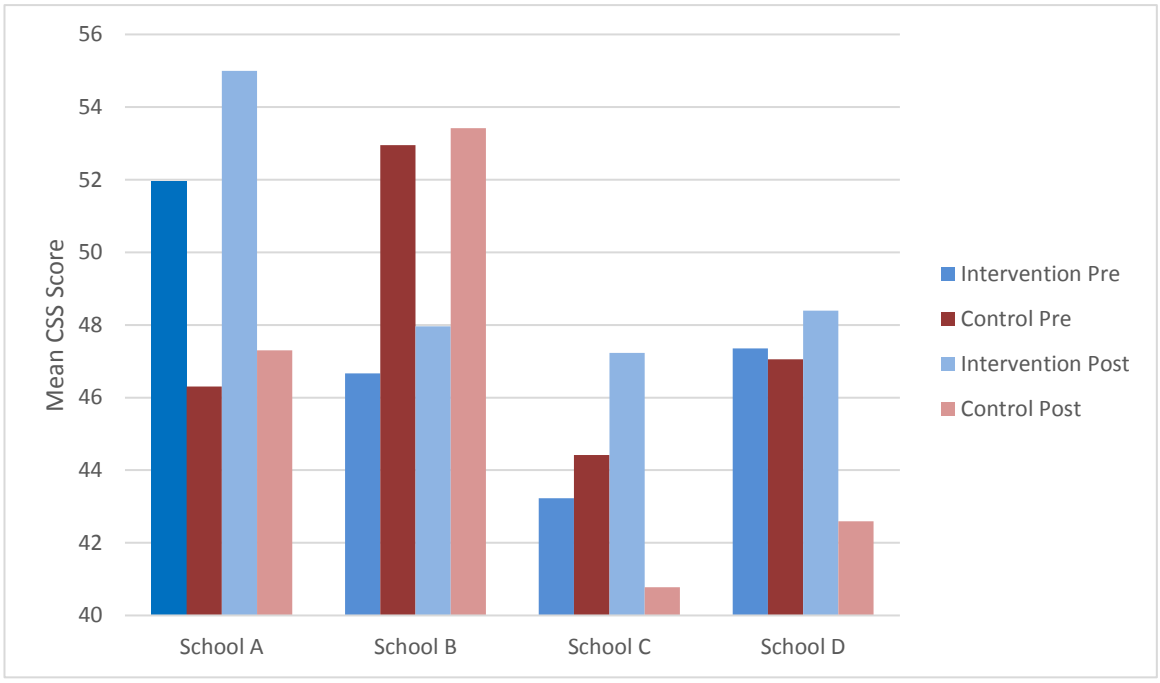
The increase for the intervention group was in the direction expected by the experimental hypothesis.

An independent samples t test found a statistically significant difference in mean change scores for intervention pupils (M = 2.23, SD = 8.48) and control pupils (M = -1.49, SD = 9.12; t (168) = 2.75, p = .007, two tailed), with the intervention group making statistically significantly greater gains. This represented a small to medium effect (d = .42).

3.9.1 Individual Schools

To understand any main effects further, the data from each individual school was scrutinised.

Figure 3.2 A graph displaying mean CSS scores at pre and post-test for intervention and control conditions in each school



As seen in Figure 3.2, the mean CSS score increased from pre to post-test in each intervention class. This was in the direction expected by the experimental hypothesis.

In schools A and B, the control class mean CSS score also increased. However, the increases in mean scores for the intervention groups were always greater than for their respective control groups.

In school C and D, the control class mean CSS score decreased. In school D, the control class decrease was greater than the intervention class increase.

There were significant differences between intervention and control classes at pre-test in:

school A t (48) = 2.18, p = .034, two tailed

school B t (41) = -2.32, p = .025, two tailed

Independent sample t tests were conducted for each school. Key data, including means, standard deviations and change scores can be found in Appendix U. Data was normally distributed, with the exception of school B which did not meet the homogeneity of variances assumption; this was accounted for within the t test.

Table 3.5 Perceptions of peer support: Outcomes of analyses for individual schools (CSS)

School	Results from statistical analyses	Cohen's *d*
School A	t (48) = .89, p = .380, two tailed	.26 (small to medium)
School B	t (29.94) = .29, p = .778, two tailed	.10 (small)
School C	t (30) = 2.98, p = .006, two tailed	1.1 (large)
School D	t (43) = 1.80, p = .079, two tailed	.55 (medium to large)

As seen in Table 3.5**,** in school C, there was a statistically significant difference in mean change scores for intervention pupils (M = 4.00, SD = 6.26) and control pupils (M = -3.63, SD = 7.62; t (30) = 2.98, p = .006, two tailed), with the intervention group making statistically significantly greater gains. This represented a large effect (d = 1.1).

No statistically significant differences in mean change scores between intervention and control classes were found in the other schools, although there was a medium to large effect found in school D.

3.9.2 A Summary of Findings

When considering the data from all schools combined, the intervention condition had statistically significantly greater gains in perceptions of peer support than the control condition (p = .007, d = .42). The null hypothesis was rejected.

There were increases in perceptions of peer support in the intervention group in all individual schools, with effect sizes in individual schools ranging from small (d = .10) to large (d = 1.1). Statistical significance was achieved in school C.

3.10 Perceptions of Peer Support: A Differential Impact (2i)

Research Question 2: Does the impact of CL on pupils' perceptions of peer support vary, depending on their initial levels of a) peer acceptance or b) peer support?

Experimental Hypothesis: There will be statistically significant differences in the increases of pupils' perceptions of peer support, according to their initial levels of a) peer acceptance and b) peer support.

Changes in CSS scores were analysed according to initial peer acceptance on Frederickson & Furnham's (1998) four contexts (WSG, WDG, PSG, and PDG).

3.10.1 Perceptions of Peer Support and Initial Levels of Peer Acceptance (2i-a)

2i- a) Perceptions of peer support and initial levels of *peer acceptance*:

Data was normally distributed, with the exception of PDG data. A one way between groups ANOVA was used to analyse WSG, WDG, and PSG data. As PDG data was not normally distributed, the Kruskal-Wallis test was used to analyse the PDG data, with median scores reported.

Analyses found no statistically significant differences between increases in perceptions of peer support for pupils who were classed as low, medium and high initial peer acceptance when defined by any of the four contexts (Table 3.6).

3.10.2 Perceptions of Peer Support and Initial Levels of Perceived Peer Support (2i-b)

2i- b) Perceptions of peer support and initial levels of *perceived peer support:*

Data was normally distributed.

A one way between groups ANOVA found statistically significant differences between increases in perceptions of peer support for pupils who were classed as low, medium and high initial perception of peer support (F (2, 80) = 6.15, p = .003) which represented a moderate to large difference (Eta Squared = .13).

Post hoc analysis using Games-Howell indicated that the significant differences were between the increases in the low group when compared to the medium and high groups. That is, the pupils who started off with lower perceptions of peer support made statistically significantly greater gains than those with medium and high initial perceptions of peer support.

Please refer to Table 3.7.

Table 3.6 Changes in perceptions of peer support (CSS) in the intervention group according to initial peer acceptance level

Initial Peer Acceptance Context	Initial Peer Acceptance Level	*N*	Mean Change Score (*median for PDG)*	SD *(range for PDG)*	One way between groups ANOVA (F) or Kruskal-Wallis (H)	Games-Howell
WSG	Low	17	1.76	7.71	*F* (2, 80) = 0.172, *p* = .842	
	Medium	46	2.72	8.66	Eta Squared = .00	
	High	20	1.50	9.02		
WDG	Low	20	1.95	7.61	*F* (2, 80) = 0.01, *p* = .986	
	Medium	44	2.34	9.63	Eta Squared = .00	
	High	19	2.26	6.68		
PSG	Low	19	2.26	7.32	*F* (2, 80) = 0.136, *p* = .873	
	Medium	46	1.87	7.89	Eta Squared = .00	
	High	18	3.11	11.14		
PDG	Low	20	2.00	29.00	*H* (2) = 0.030, *p* = .985	
	Medium	45	1.50	54.00		
	High	18	2.00	29.00		

Table 3.7 Changes in perceptions of peer support (CSS) in the intervention group according to initial level of perceived peer support

	N	Mean Change Score	*SD*	One way between groups ANOVA	Games-Howell
Low	20	7.35	7.51	F (2, 80) = 6.15, p = .003	Low – medium (p = .016)
Medium	44	1.41	7.71	Eta Squared = .13	Low – high (p = .008)
High	19	-1.26	9.07		

3.10.3 A Summary of Findings

There were no statistically significant differences in increases in perceptions of peer support for children with initially low, medium or high peer acceptance. That is, the degree to which CL had an impact did not appear to be influenced by initial levels of peer acceptance. The null hypothesis was rejected.

However, pupils who had lower initial levels of perceptions of peer support made statistically significantly greater gains in perceptions of peer support than either their peers with initially medium or high levels (p = .003**,** Eta Squared = .13). The null hypothesis was rejected.

3.11 Perceptions of Peer Support: A Differential Impact (2ii)

2ii

Experimental Hypothesis: There will be statistically significantly greater increases in perceptions of peer support for pupils with low initial perceptions of peer support in the experimental than control group.

Pre and post-test median data for intervention and control conditions is presented in Table 3.8.

2ii Low perceptions of peer support (intervention v control):

Data was not normally distributed, so non-parametric tests were used to analyse the data, with median scores reported.

The Mann-Whitney U Test indicated that intervention and control groups were not statistically significantly different at pre-test (U = 193.00, z = 0.45, p = .652), nor at post-test (U = 188.50, z = -0.56, p = .574).

The Wilcoxon Signed Rank Test indicated that low pupils in the intervention group made statistically significant increases pre to post-test (T = 193.00, z = 3.29, p = .001). This represented a large effect (r = .52). No statistically significant changes were found in the low control group (T = 132.50, z = 1.51, p = .131).

Table 3.8 Changes in perceptions of peer support for pupils in intervention and control conditions with low initial levels of perceived peer support

		N	Median	Range
Intervention	Pre Test	20	36.00	24.0
	Post Test	20	41.00	45.00
Control	Pre Test	21	37.00	17.00
	Post Test	21	39.00	39.00

3.11.1 A Summary of Findings

Children with initially low perceptions of peer support made statistically significant increases in perceptions of peer support in the intervention condition (p = .001, r = .52) but not in the control condition. The null hypothesis was rejected.

3.12 Peer Acceptance

Peer acceptance was measured using the Social Inclusion Survey (Frederickson and Furnham, 1998).

***Research Question 3*:** What is the impact of CL on pupils' peer acceptance?

***Experimental Hypothesis**:* There will be a statistically significantly greater increase in pupils' peer acceptance following CL (experimental condition), when compared to pupils in a control condition.

Results are presented in four sections, according to the four peer acceptance contexts (WSG, WDG, PSG, and PDG). Within each section, overall findings are presented, followed by a closer look at individual schools.

3.12.1 Changes in Work Same Gender Peer Acceptance

In the main analysis, data from all four schools was combined (Table 3.9). Data was normally distributed.

Table 3.9 Work same gender peer acceptance: Pre and post-test overall means and change scores for intervention and control conditions (WSG)

		N	Mean	*SD*	Change Score	*SD*
Intervention	Pre Test	101	45.58	35.48		
					5.54	25.70
	Post Test	101	51.13	37.95		
Control	Pre Test	102	49.93	39.25		
					-7.97	27.44
	Post Test	102	41.97	49.45		

Intervention and control group pre-test mean scores were broadly similar. An independent samples t test indicated that they were not statistically significantly different (t (201) = -0.83, p = 0.409).

Mean scores increased for the intervention group from pre-test to post-test, and decreased for the control group over the same period. This is illustrated in Figure 3.3.

Figure 3.3 A graph displaying overall mean WSG scores at pre and post-test for intervention and control conditions

The increase for the intervention group was in the direction expected by the experimental hypothesis.

An independent samples t test found that there was a statistically significant difference in change scores for intervention pupils (*M* = 5.54, *SD* = 25.70) and control pupils (*M* = -7.97, *SD* = 27.44; *t* (201) = 3.62, *p* = .000, two tailed). This represented a medium to large effect (*d* = .51).

3.12.1.1 Individual Schools

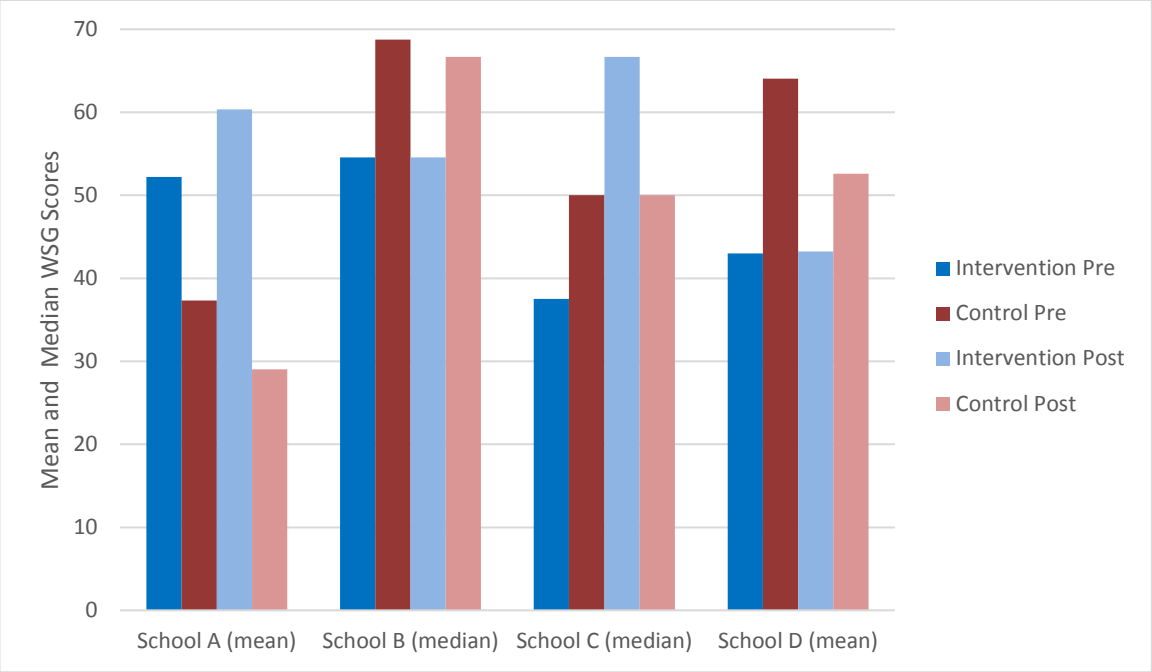
To understand any main effects further, the data from each individual school was scrutinised.

Schools A and D were normally distributed, so mean data has been presented. Independent samples t tests were used to test for significance. As individual school data for schools B and C were not normally distributed, median rather than mean data has been presented for those schools. Mann-Whitney U and Wilcoxon Signed Rank tests were used to test for significance.

As illustrated in Figure 3.4, the mean or median WSG scores increased from pre to post-test in three intervention classes. This was in the direction expected by the experimental hypothesis. The median was stable in school B from pre to post-test.

In schools A, B and D, the control class mean or median decreased from pre to post-test. In school C, the control class median remained stable from pre to post-test.

Figure 3.4 A graph displaying mean and median WSG scores at pre and post-test for intervention and control conditions in each school



Key data including means and medians, and change scores can be found in Appendix V.

Table 3.10 provides a summary of statistical analyses for each individual school.

Table 3.10. Work same gender peer acceptance: Outcomes of analyses for individual schools (WSG)

School	Results from statistical analyses	Cohen's *d* or Pearson's *r*
School A	*t* (53) = 2.54, *p* = .014, two tailed	*d* = .70 (medium to large)
School B	Wilcoxon Signed Ranks, pre to post Intervention: *T* = 91.00, *z* = -0.86, *p* = .392) Control: *T* = 110.00, *z* = -0.19, *p* = .848) Mann Whitney U, intervention and control at Pre test: *U* = 337.00, *z* = 0.23, *p* = .821 Post test: *U* = 380.00, *z* = 1.04, *p* = .299	*r* = .12 (small to medium)
School C	Wilcoxon Signed Ranks, pre to post Intervention: T = 113.00, z = 3.024, p = 0.002 Control: T = 47.50, z = -0.713, p = 0.476 Mann Whitney U, intervention and control at Pre test: U = 219.50, z = 0.597, p = 0.602 Post test: U = 140.00, z = -1.626, p = 0.108	*r* = .48 (medium to large)
School D	*t* (55) = 1.71, *p* = .093, two tailed	d = .45 (small to medium)

As seen in Table 3.10, in school A, there was a statistically significant difference in mean change scores for intervention pupils (*M* = 8.15, *SD* = 23.95) and control pupils (*M* = -8.28, *SD* = 24.12; *t* (53) = 2.54, *p* = .014, two tailed), with the intervention group making statistically significantly greater gains. This represented a medium to large effect (*d* = .70).

In school C, the intervention group made statistically significant gains from pre to post-test (T = 113.00, z = 3.024, p = .002). This represented a medium to large effect (*r* = .48). No statistically significant gains were found for the control group.

No statistically significant differences were found in schools B or D, with the exception of statistically significantly differences between the intervention and control class at pre-test in school D (*t* (55) = -2.68, *p* = .010). Effect sizes were small to medium.

3.12.2 Changes in Work Different Gender Peer Acceptance

In the main analysis, data from all four schools was combined (Table 3.11). Data was normally distributed.

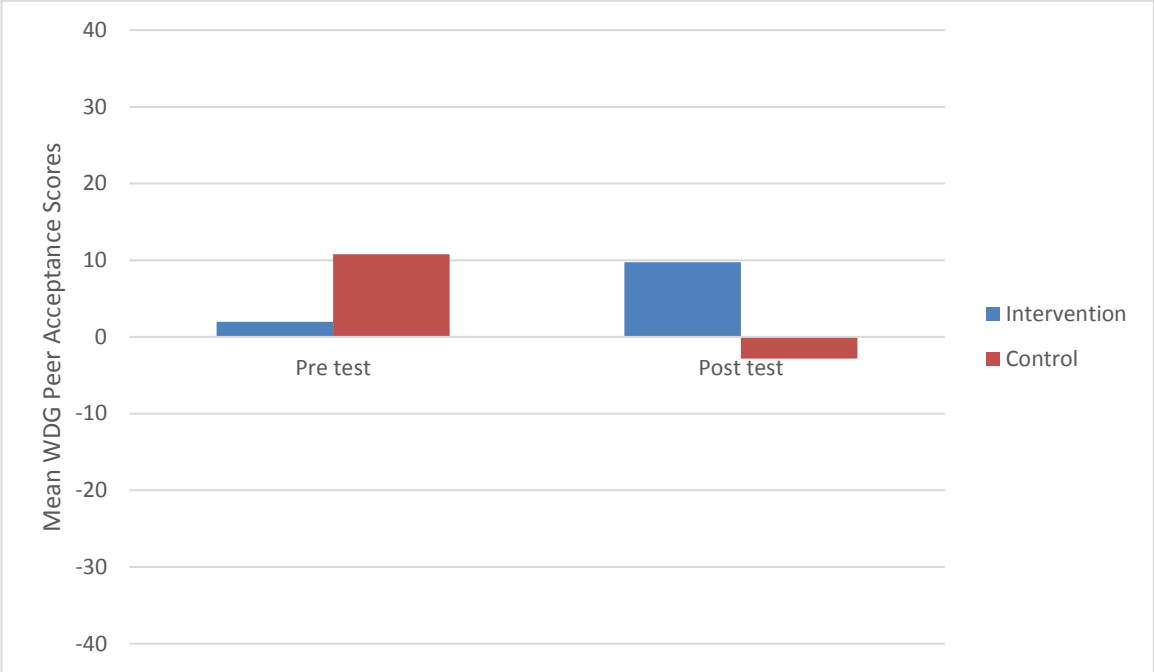
Table 3.11. Work different gender peer acceptance: Pre and post-test overall means and change scores for control and intervention conditions (WDG)

		N	Mean	*SD*	Change Score	*SD*
Intervention	Pre Test	101	1.97	40.78		
					7.77	30.22
	Post Test	101	9.72	37.85		
Control	Pre Test	102	10.79	37.79		
					-13.62	27.37
	Post Test	102	-2.82	40.03		

Intervention and control group means were quite different at pre-test, with the intervention group mean lower than that of the control group. However, an independent samples t test indicated that they were not statistically significantly different (t (201) = -1.60, p = 0.111).

Mean scores increased for the intervention group from pre to post-test, and decreased for the control group over the same period. At post-test, the intervention group mean score was higher than that of the control group. This is illustrated in Figure 3.5.

Figure 3.5 A graph displaying overall mean WDG scores at pre and post-test for intervention and control conditions



The increase for the intervention group was in the direction expected by the experimental hypothesis.

An independent samples t test found that there was a statistically significant difference in change scores for intervention pupils (*M* = 7.77, *SD* = 30.22) and control pupils (*M* = -13.62, *SD* = 27.37; *t* (201) = 5.28, p = .000, two tailed). This represented a medium to large effect (*d* = .74).

3.12.2.1 Individual Schools

To understand any main effects further, the data from each individual school was scrutinised. Data was normally distributed.

As illustrated in Figure 3.6, the intervention class mean WDG scores increased from pre to post-test in schools A, C and D. This was in the direction expected by the experimental hypothesis. However, there was an unexpected decrease in WDG scores from pre to post-test in school B.

In all four schools, the control class mean work different gender score decreased from pre to post-test. In school B, the decrease was less for the intervention class than it was for the control class.

Figure 3.6 A graph displaying mean WDG scores at pre and post-test for intervention and control conditions in each school

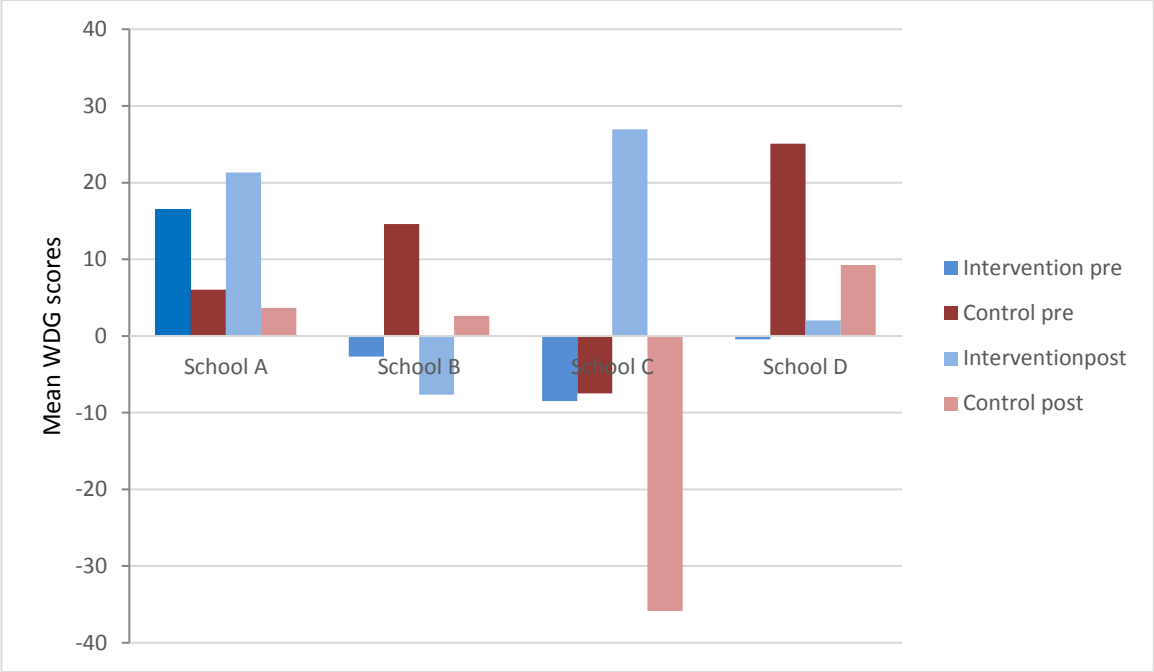


Table 3.12 provides a summary of results from independent sample t tests conducted for each school. Key data, including means, standard deviations and change scores can be found in Appendix W.

Table 3.12 Work different gender peer acceptance: Outcomes of analyses for individual schools (WDG)

School	Results from statistical analyses	Cohen's *d*
School A	*t* (53) = 1.08, *p* = .284, two tailed	.30 (small to medium)
School B	*t* (49) = 1.06, *p* = .295, two tailed	.30 (small to medium)
School C	*t* (38) = 6.44, *p* = .000, two tailed	2.04 (large)
School D	*t* (55) = 2.49, *p* = .016, two tailed	.67 (medium to large)

As seen in Table 3.12, in school C, there was a statistically significant difference in mean change scores for intervention pupils (*M* = 35.42, *SD* = 36.10) and control pupils (*M* = -28.35, *SD* = 25.65; *t* (38) = 6.44, *p* = .000, two tailed), with the intervention group making statistically significantly greater gains. This represented a large effect (*d* = 2.04).

In school D, there was also a statistically significant difference in mean change scores for intervention pupils (*M* = 2.52, *SD* = 28.61) and control pupils (*M* = -15.82, *SD* = 26.87; *t* (55) = 2.49, *p* = .016, two tailed, with the intervention group making statistically significantly greater gains. This represented a medium to large effect (*d* = .67). However, it should be noted that the intervention class was statistically significantly lower than the control class at pre-test (*t* (55) = -2.99, *p* = .004).

No statistically significant differences were found in schools A or B. Effect sizes were small to medium.

3.12.3 Changes in Play Same Gender Peer Acceptance

In the main analysis, data from all four schools was combined (Table 3.13). Data was normally distributed.

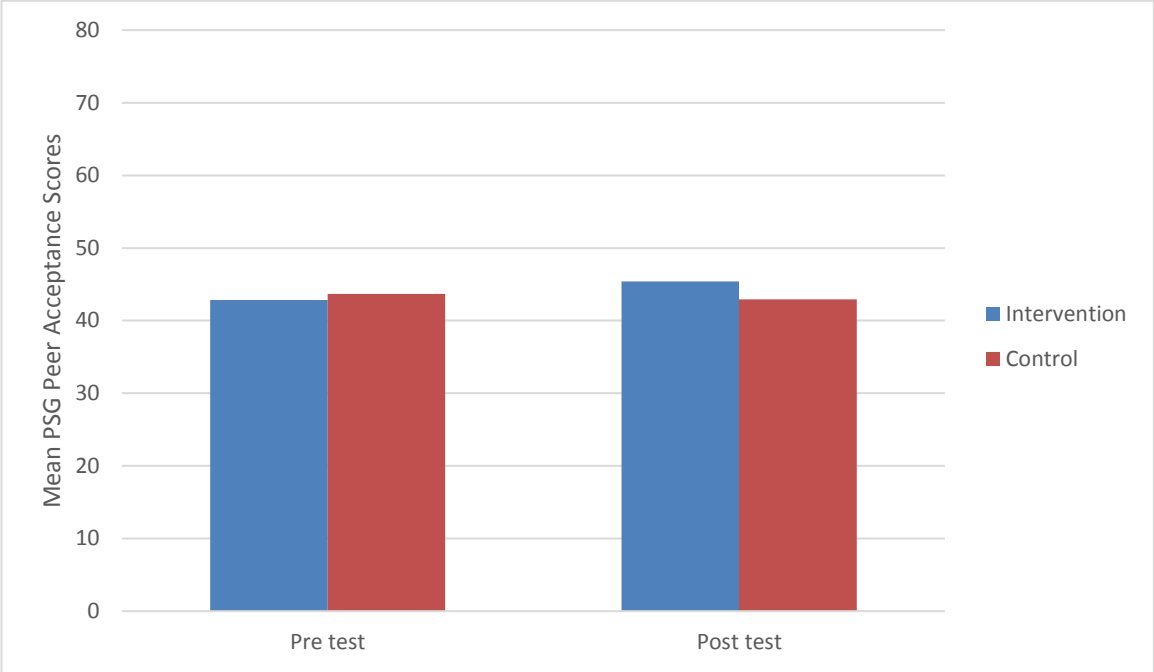
Table 3.13 Play same gender peer acceptance: Pre and post-test overall means and change scores for intervention and control conditions (PSG)

		N	Mean	*SD*	Change Score	*SD*
Intervention	Pre Test	101	42.83	35.52		
					2.56	27.92
	Post Test	101	45.40	39.54		
Control	Pre Test	102	43.69	40.29		
					-0.76	27.97
	Post Test	102	42.95	49.44		

Intervention and control group pre-test mean scores were broadly similar. An independent samples t test indicated that these were not statistically significantly different (t (201) = -0.16, p = 0.872).

Mean scores increased for the intervention group from pre to post-test, and decreased for the control group over the same period. This is illustrated in Figure 3.7.

Figure 3.7 A graph displaying overall mean PSG scores at pre and post-test for intervention and control conditions



The increase for the intervention group was in the direction expected by the experimental hypothesis.

An independent samples t test found that there was no statistically significant difference in change scores for intervention pupils (*M* = 2.56, *SD* = 27.92) and control pupils (*M* = -0.76, *SD* = 27.97; *t* (201) = .848, *p* = .398, two tailed). This represented a small to medium effect (*d* = .12).

Additional analysis without school B, indicated statistically significant differences in change scores for intervention (*M* = 7.70, *SD* = 25.99) and control pupils (*M* = -1.42, *SD* = 24.05; *t* (150) = 2.25, *p* = .026), representing a small to medium effect (*d* = .37). The reasons for this additional analysis are discussed further in section 4.1.5 and 4.1.9.

3.12.3.1 Individual Schools

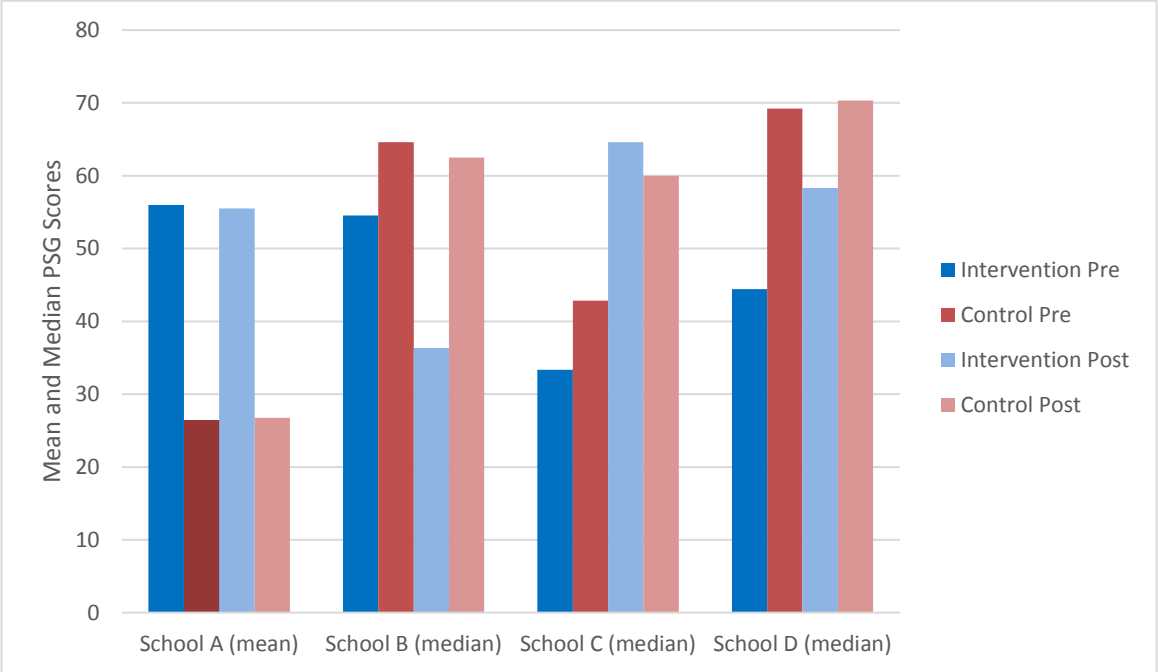
To understand any main effects further, the data from each individual school was scrutinised.

School A was normally distributed, so mean data has been presented. An independent samples t test was used to test for significance. As individual school data for schools B, C and D were not normally distributed, median data has been presented. Mann-Whitney U and Wilcoxon Signed Rank tests were used to test for significance.

As illustrated in Figure 3.8, the median score increased from pre to post-test in schools C and D. This was in the direction expected by the experimental hypothesis. However, there was a slight decrease from pre to post-test in the mean score in school A, and a larger decrease in median score in school B.

In schools A and B, the control class mean or median decreased from pre to post-test. However, in schools C and D, there was an increase in median score over the same period.

Figure 3.8 A graph displaying mean and median PSG scores at pre and post-test for intervention and control conditions in each school



Key data including means and medians, and change scores can be found in Appendix X.

Table 3.14 provides a summary of statistical analyses for each individual school.

Table 3.14. Play same gender peer acceptance: Outcomes of analyses for individual schools (PSG)

School	Results from statistical analyses	Cohen's *d* or Pearson's *r*
School A	*t* (53) = -0.13, *p* = 0.899, two tailed).	*d*= .03 (small)
School B	Wilcoxon Signed Ranks, pre to post Intervention: *T* = 61.00, *z* = -2.35, *p* = .019 Control: *T* = 164.50, *z* = -0.42, *p* = 0.678) Mann Whitney U, intervention and control at Pre test: *U* = 325.00, *z* = .00, *p* = 1.00 Post test: *U* = 409.00, *z* = 1.59, *p* = .113	*r* = .33 (medium to large)
School C	Wilcoxon Signed Ranks, pre to post Intervention: *T* = 141.00, *z* = 2.423, *p* = .015 Control: *T* = 61.00, *z* = 0.06, *p* = .955 Mann Whitney U, intervention and control at Pre test: *U* = 258.00, *z* = 1.57, *p* = .121 Post test: *U* = 186.50, *z* = -0.37, *p* = .718	*r* = .38 (medium to large)

School D	Wilcoxon Signed Ranks, pre to post Intervention: T = 242.50, z = 1.70, p = .088 Control: T = 116.00, z = -0.97, p = .331 Mann Whitney U, intervention and control at Pre test: U = 567.50, z = 2.58, p = .010 Post test: U = 465.50, z = .950, p = .342	r = .23 (small to medium)

As seen in Table 3.14, in school B, there was a statistically significant change in the scores of the intervention group from pre to post-test (T = 61.00, z = -2.35, p = .019) which represented a medium to large effect (r = .33). However, as indicated in Figure 3.8, this was a significant *decrease* in scores, and therefore not in the direction of the experimental hypothesis.

On the other hand, in school C, there was a statistically significant change in the scores of the intervention group from pre to post-test (T = 141.00, z = 2.42, p = .015) which represented a medium to large effect (r = .38). This was a significant increase in scores, in the direction of the experimental hypothesis.

In school D, intervention and control groups were significantly different at pre-test, but not at post-test. However, there were no statistically significant changes in either the intervention (p = .088) or control (p = .331) group pre to post-test.

No statistically significant differences were found in school A, with the exception of statistically significant differences between intervention and control classes at pre-test (t (53) = 3.16, p = .003).

3.12.4 Changes in Play Different Gender Peer Acceptance

In the main analysis, data from all four schools was combined (Table 3.15). Data was normally distributed.

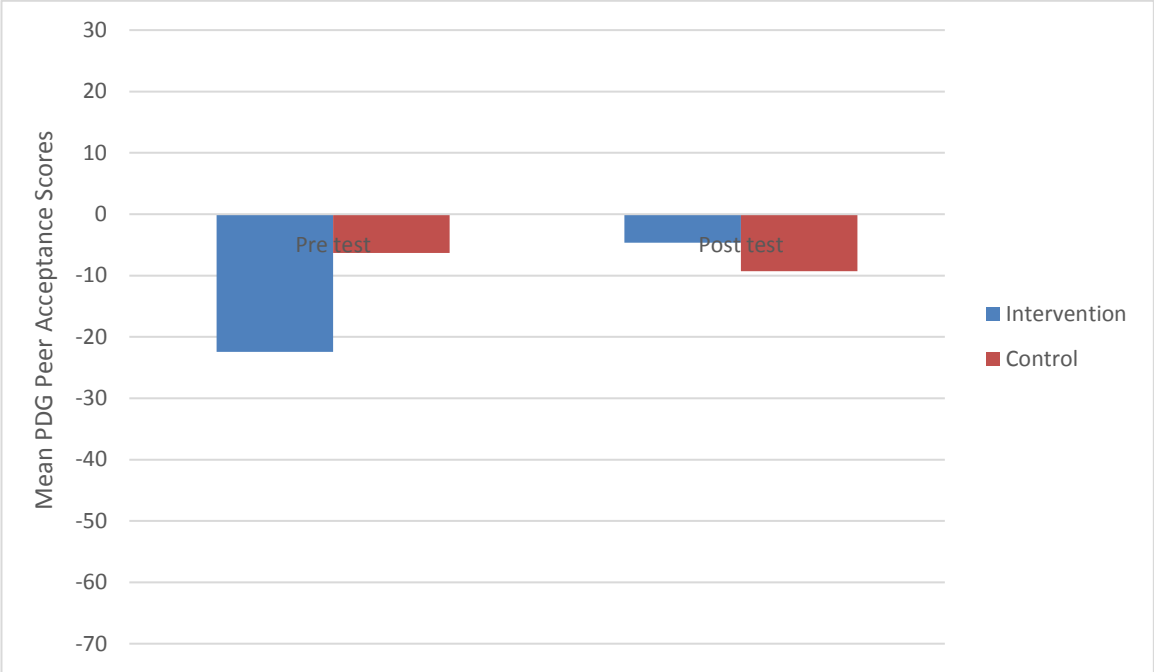
Table 3.15 Play different gender peer acceptance: Pre and post-test overall means and change scores for intervention and control conditions (PDG)

		N	Mean	*SD*	Change Score	*SD*
Intervention	Pre Test	101	-22.42	37.85		
					17.72	33.59
	Post Test	101	-4.67	38.43		
Control	Pre Test	102	-6.29	40.19		
					-2.98	25.46
	Post Test	102	-9.29	46.59		

Intervention and control group means were quite different at pre-test, with an independent samples t test indicating that the intervention group mean was statistically significantly lower than that of the control group at pre-test (t (201) = -2.94, p = .004)

Mean scores increased for the intervention group from pre to post-test, and decreased for the control group over the same period. This is illustrated in Figure 3.9.

Figure 3.9 A graph displaying overall mean PDG scores at pre and post-test for intervention and control conditions



The increase for the intervention group was in the direction expected by the experimental hypothesis.

An independent samples t test found that there was a statistically significant difference in change scores for intervention pupils (*M* = 17.72, *SD* = 33.59) and control pupils (*M* = -2.98, *SD* = 25.46; *t* (201) = 4.95, *p* = .000, two tailed). This represented a medium to large effect (*d* = .70).

3.12.4.1 Individual schools

To understand any main effects further, the data from each individual school was scrutinised.

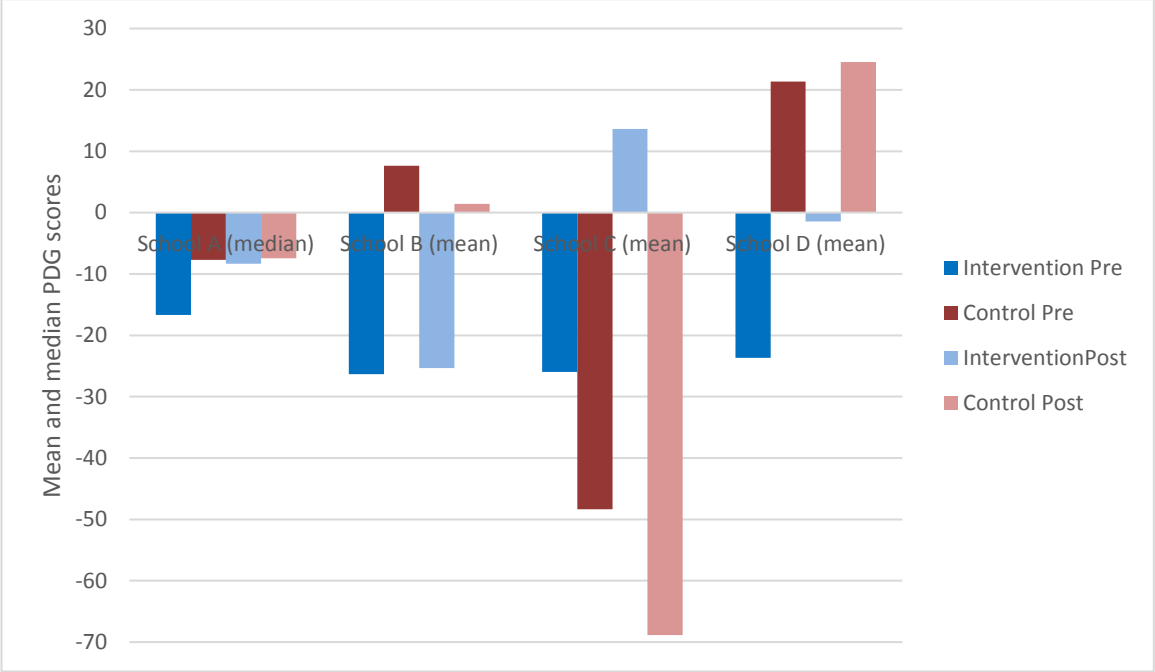
As individual school data for school A was not normally distributed, median rather than mean data has been presented. Mann Whitney U and Wilcoxon Signed Rank tests were used to test for significance. The remaining schools

were normally distributed, so mean data has been presented. Independent samples t tests were used to test for significance.

As illustrated in Figure 3.10, the mean or median scores increased from pre to post-test in each intervention class. This was in the direction expected by the experimental hypothesis.

In schools A, B and C, the control class median or means decreased from pre to post-test. In school D, the control class mean increased from pre to post-test.

Figure 3.10 A graph displaying mean and median PDG scores at pre and post-test for intervention and control conditions in each school



Key data including means and medians, and change scores can be found in Appendix Y.

Table 3.16 provides a summary of statistical analyses for each individual school.

Table 3.16. Play different gender peer acceptance: Outcomes of analyses for individual schools (PDG)

School	Results from statistical analyses	Cohen's *d* or Pearson's *r*
School A	Wilcoxon Signed Ranks, pre to post Intervention: *T* = 152.50, *z* = 2.32, *p* = .020 Control: *T* = 247.00, *z* = 1.395, *p* = .163 Mann Whitney U, intervention and control at Pre test: *U* = 402.50, *z* = 0.413, *p* = .679 Post test: *U* = 355.00, *z* = -0.39, *p* = .698)	*r* = .32 (medium to large)
School B	*t* (49) = 1.08, *p* = 0.284, two tailed)	*d* = .31 (small to medium)
School C	*t* (38) = 6.05, *p* = .000, two tailed	*d* = 1.96 (large)
School D	*t* (55) = 2.16, *p* = .035, two tailed	*d* = .58 (medium to large)

As seen in Table 3.16, in school C, there were statistically significant differences in mean change scores for intervention pupils (*M* = 39.58*, SD* = 35.99*)* and control pupils (*M* = -20.51*, SD* = 26.09; *t* (38) = 6.05, *p* = .000, two tailed), with the intervention group making statistically significantly greater gains. This represented a large effect (*d* = 1.96).

Similarly, in school D, there were statistically significant differences in mean change scores for intervention pupils (*M* = 22.10*, SD* = 39.13*)* and control pupils (*M* = 3.29, *SD = 24.77; t* (55) = 2.16, *p* = .035, two tailed), with the intervention group making statistically significantly greater gains. This represented a medium to large effect (*d* = .58). However, it should be noted that the intervention class was statistically significantly lower than the control class at pre-test (*t* (55) = -5.26, *p* = .000).

In school A, the intervention group made statistically significant gains from pre to post-test (*T* = 152.50, *z* = 2.32, *p* = .020). This represented a medium to large effect (*r* = .32). No statistically significant gains were found for the control group.

No statistically significant differences were found in school B, with the exception of statistically significant differences between intervention and control classes at pre-test (*t* (49) = -2.95, *p* = .005). The effect size was small to medium.

3.12.5 A Summary of Findings

When considering the data from all schools combined, the null hypothesis could be rejected for three of the four social contexts. The intervention group made statistically significantly greater gains than the control group with regards to WSG (*p* = .000, *d* = .51), WDG (*p* = .000, *d* = .74), and PDG (*p* = .000, *d* = .70). These represented medium to large and large effects. Significant differences at pre-test between intervention and control conditions suggest some caution when interpreting the findings of PDG.

With regards to individual school level changes, findings were predominantly in favour of the intervention classes, as follows:

WSG: Three of the intervention classes increased pre to post-test, and one remained stable (school B). Effect sizes for individual schools ranged from small to medium (r = .12) to medium to large (d = .70). Statistical significance was achieved in school A and C.

WDG: Three of the intervention classes increased pre to post-test, and one (school B) decreased. Statistical significance was achieved in school C and D, with effect sizes across the four schools ranging from small to medium (d = .30) to large (d = 2.04).

PDG: All four intervention classes increased from pre to post-test. Statistical significance was found in school A (r = .32, medium to large effect), school C (d = 1.96, large effect), and school D (d = .58, medium to large effect), although in school D, the intervention group was significantly lower than the control group at pre-test.

The exceptions appeared to be in school B, and this is discussed further in section 4.1.9.

In the fourth context, PSG, when all schools were combined, although the intervention group increased in line with the experimental hypothesis, this did not reach statistical significance and the effect size was weak (p = .398, d = .12). The null hypothesis cannot be rejected for this social context.

Two of the intervention classes increased pre to post-test with regards to PSG peer acceptance, with this reaching statistical significance in school C. In the other two intervention classes, there was a decrease pre to post-test, which was statistically significant in school B, and not expected by the experimental hypothesis.

When school B was removed from the overall analysis on PSG, there was a statistically significant difference between the changes in scores for intervention and control groups, which represented a small to medium effect (p = .026, d = .37).

3.13 Peer Acceptance: A Differential Impact (4i)

Research Question 4: Does the impact of CL on pupils' peer acceptance vary, depending on pupils' initial levels of peer acceptance?

4i

Experimental Hypothesis*:* There will be statistically significant differences in the increases of pupils' peer acceptance, according to their initial levels of peer acceptance.

Results are presented in four sections, according to the four peer acceptance contexts (WSG, WDG, PSG, and PDG). Within each section, analyses have been carried out according to initial peer acceptance on each of the four contexts.

Data has been analysed for the intervention group only, with initial peer acceptance scores used to categorise pupils as low, medium or high.

3.13.1 Work Same Gender

Key descriptive data can be found in Table 3.17.

Data was normally distributed, with the exception of PDG data. A one way between groups ANOVA was used to analyse WSG, WDG, and PSG data. The Kruskal-Wallis test was used to analyse the PDG data, with median scores reported.

Table 3.17 Changes in work same gender peer acceptance in the intervention group for pupils with initially low, medium, and high peer acceptance (descriptive and inferential data)

Initial Peer Acceptance Context	Initial Peer Acceptance Level	*N*	Mean Change Score (*median for PDG)*	*SD (range for PDG)*	One way between groups ANOVA (F) or Kruskal-Wallis (H)	Games-Howell
WSG	Low	24	12.01	21.42	*F* (2, 98) = 4.30, *p* = .016	Low-high (*p* = .006)
	Medium	53	9.18	27.10	Eta Squared = .08	Medium-high (*p* = .010)
	High	24	-5.96	16.60		
WDG	Low	23	10.55	23.76	*F* (2, 98) = 0.58, *p* = .560	
	Medium	53	5.96	26.53	Eta Squared = .01	
	High	25	2.93	20.60		
PSG	Low	25	6.13	20.66	*F* (2, 98) = 1.90, *p* = .156	
	Medium	51	10.09	26.57	Eta Squared = .04	
	High	25	-1.44	22.47		

	Low	25	9.09	97.67	H (2) = 0.48, *p* = .788
PDG	Medium	54	2.00	121.21	
	High	22	0.00	85.61	

Based on trends, the greatest increases in WSG peer acceptance appeared to be for children rated as low peer accepted on the WSG, WDG and PDG, and children rated as medium peer accepted on the PSG.

There were statistically significant differences found in increases in work same gender peer acceptance scores for pupils who were classed as low, medium and high initial peer acceptance when defined using their WSG score (F (2, 98) = 4.30, p = .016). This represented a moderate to large difference (Eta Squared = .08).

Post hoc analysis using Games-Howell indicated that both the low and medium groups gained statistically significantly more than the high group.

3.13.2 Work Different Gender

Key descriptive data can be found in Table 3.18.

Data was normally distributed, with the exception of PDG data. A one way between groups ANOVA was used to analyse WSG, WDG, and PSG data. The Kruskal-Wallis test was used to analyse the PDG data, with median scores reported.

Table 3.18. Changes in work different gender peer acceptance in the intervention group for pupils with initially low, medium, and high peer acceptance (descriptive and inferential data)

Initial Peer Acceptance Context	Initial Peer Acceptance Level	*N*	Mean Change Score (*median for PDG)*	*SD* (*median for PDG)*	One way between groups ANOVA (F) or Kruskal-Wallis (H)	Games-Howell
WSG	Low	24	9.74	29.13	*F* (2, 98) = 0.56, *p* = .575	
	Medium	53	9.45	30.01	Eta Squared = .01	
	High	24	2.06	32.29		
WDG	Low	23	20.28	32.17	*F* (2, 98), = 7.89, *p* = .001	Low-high (*p* = .001)
	Medium	53	11.03	29.26	Eta Squared = .14	Medium – high (*p* = .002)
	High	25	-10.67	21.80		
PSG	Low	25	14.20	32.51	*F* (2, 98) = 1.07, *p* = .346	
	Medium	51	3.71	29.07	Eta Squared = .02	
	High	25	9.62	30.11		

	Low	25	16.00	133.33	H (2) = 4.65, p = 0.098
PDG	Medium	54	8.34	120.45	
	High	22	0.00	98.45	

Based on trends, the greatest increases in WDG peer acceptance appeared to be for children rated as low peer accepted in all four contexts.

There was a statistically significant difference between increases in work different gender peer acceptance scores for pupils who were classed as low, medium, and high initial peer acceptance when defined using their WDG score (F (2, 98), = 7.89, p = .001). This represented a large difference (Eta Squared = .14).

Post hoc analysis using Games-Howell indicated that both the low and medium groups gained statistically significantly more than the high group.

3.13.3 Play Same Gender

Key descriptive data can be found in Table 3.19.

Data was normally distributed, with the exception of PDG data. A one way between groups ANOVA was used to analyse WSG, WDG, and PSG data. The Kruskal-Wallis test was used to analyse the PDG data, with median scores reported.

Table 3.19. Changes in play same gender peer acceptance in the intervention group for pupils with initially low, medium, and high peer acceptance (descriptive and inferential data)

Initial Peer Acceptance Context	Initial Peer Acceptance Level	*N*	Mean Change Score (*median for PDG)*	*SD (range for PDG)*	One way between groups ANOVA (F) or Kruskal-Wallis (H)	Games-Howell
WSG	Low	24	4.60	28.11	*F* (2, 98) = 0.10, *p* = .907	
	Medium	53	2.31	28.52	Eta Squared = .00	
	High	24	1.09	27.46		
WDG	Low	23	-0.36	30.87	*F* (2, 98) = 0.66, *p* = .521	
	Medium	53	5.59	26.48	Eta Squared = .01	
	High	25	-1.17	28.51		
PSG	Low	25	8.92	32.86	*F* (2, 98) = 5.90, *p* = .004	Low – high (*p* = .024)
	Medium	51	7.21	24.53	Eta Squared = .11	Medium – high (*p* = .003)
	High	25	-13.27	23.80		

	Low	25	0.00	104.54	H (2) = 1.83, *p* = .402
PDG	Medium	54	0.00	139.40	
	High	22	9.10	104.55	

Based on trends, the greatest increases in PSG peer acceptance appeared to be for children rated as low peer acceptance on the WSG and PSG, but medium group as rated by WDG, and high group as rated by PDG.

There was a statistically significant difference between increases in play same gender peer acceptance levels for pupils who were classed as low, medium and high initial peer acceptance when defined using their PSG score (F (2, 98) = 5.90, p = .004). This represented a moderate to large difference (Eta Squared = .11).

Post hoc analysis using Games-Howell indicated that both the low and medium groups gained statistically significantly more than the high group.

3.13.4 Play Different Gender

Key descriptive data can be found in Table 3.20.

Data was normally distributed, with the exception of PDG data. A one way between groups ANOVA was used to analyse WSG, WDG, and PSG data. The Kruskal-Wallis test was used to analyse the PDG data, with median scores reported.

Table 3.20. Changes in play different gender peer acceptance in the intervention group for pupils with initially low, medium, and high peer acceptance (descriptive and inferential data)

Initial Peer Acceptance Context	Initial Peer Acceptance Level	*N*	Mean Change Score (*median for PDG)*	*SD (range for PDG)*	One way between groups ANOVA (F) or Kruskal-Wallis (H)	Games-Howell
WSG	Low	24	25.46	35.27	F (2, 98) = 1.59, *p* = .209	
	Medium	53	18.44	34.51	Eta Squared = .03	
	High	24	8.39	28.56		
WDG	Low	23	32.10	42.78	F (2, 98) = 5.18, *p* = .007	Low – high (*p* = .0016)
	Medium	53	18.78	30.27	Eta Squared = .10	Medium – high (*p* = .032)
	High	25	2.25	24.28		
PSG	Low	25	27.04	38.11	F (2. 98) = 1.43, *p* = .245	
	Medium	51	16.07	32.91	Eta Squared = .03	
	High	25	11.76	29.32		

	Low	25	22.22	139.34	H (2) = 12.81, *p* = .002	Low – high (*p* = .001)
PDG	Medium	54	9.09	127.89		r = 1.87
	High	22	-4.00	70.70		

Based on trends, the greatest increases in PDG peer acceptance appeared to be for children rated as low peer acceptance across all four contexts.

There was a statistically significant difference between increases in play different gender peer acceptance scores for pupils who were classed as low, medium and high initial peer acceptance when defined using their PDG score ($H(2) = 12.81$, $p = .002$). The low group gained significantly more than the high group, representing a large effect ($r = 1.87$).

There was also a statistically significant difference between increases in play different gender peer acceptance scores for pupils who were classed as low, medium and high initial peer acceptance when defined using their WDG score ($F(2, 98) = 5.18$, $p = .007$). Post hoc analysis using Games-Howell indicated that both the low and medium groups gained statistically significantly more than the high group. This represented a moderate to large difference (Eta Squared = .10).

3.13.5 A Summary of Findings

In all four contexts (WSG, WDG, PSG, and PDG), when looking within the same context, pupils with initially lower levels of peer acceptance made statistically significantly greater gains than those with higher initial levels. Often, pupils with medium levels of peer acceptance also made statistically significantly greater gains than those with higher levels, but never greater than those with low peer acceptance. Effect sizes were moderate to large, or large. The null hypothesis was rejected.

3.14 Peer Acceptance: A Differential Impact (4ii)

4ii

Experimental Hypothesis: There will be statistically significantly greater increases in perceptions of peer acceptance for pupils with low initial levels of peer acceptance in the experimental than control group.

Results are presented in four sections, according to the four peer acceptance contexts (WSG, WDG, PSG, and PDG).

Analysis was concerned with those pupils who were classed as having low initial peer acceptance (intervention v control).

3.14.1 Work Same Gender

When considering pupils who started off as low in WSG peer acceptance, the intervention and control groups were not statistically different at pre-test (t (47) = 0.54, p = .592). Data was normally distributed.

There was a statistically significant difference between the gains made by the intervention and control group (t (47) = 3.78, p = .000), with the children classed as low initial peer acceptance in the intervention group making statistically significantly greater gains than those in the control group. This represented a large effect (d = 1.08).

Table 3.21 Changes in work same gender peer acceptance for pupils in intervention and control conditions with low initial peer acceptance (WSG)

		N	*M* at pre test	*SD*	Mean Change Score	SD
Intervention (low)	Pre Test	24	-3.92	20.22	12.01	21.42
	Post Test	24	8.06	27.52		
Control (low)	Pre Test	25	-7.76	28.60	-18.23	33.12
	Post Test	25	-25.98	41.83		

3.14.2 Work Different Gender

When considering pupils who started off as low in WDG peer acceptance, the intervention and control groups were not statistically different at pre-test (t (45) = -1.97, p = .055). Data was normally distributed.

There was a statistically significant difference between the gains made by the intervention and control group (t (45) = 3.17, p = .003), with the children classed as low initial peer acceptance in the intervention group making statistically significantly greater gains than those in the control group. This represented a large effect (d = .92).

Table 3.22 Changes in work different gender peer acceptance for pupils in intervention and control conditions with low initial peer acceptance (WDG)

		N	*M* at pre test	*SD*	Mean Change Score	SD
Intervention (low)	Pre Test	23	-52.70	18.63	20.28	32.17
	Post Test	23	-32.44	30.90		
Control (low)	Pre Test	24	-41.04	21.76	-5.47	23.01
	Post Test	24	-46.51	26.28		

3.14.3 Play Same Gender

When considering pupils who started off as low in PSG peer acceptance, Mann Whitney U indicated that intervention and control groups were not statistically significantly different at pre-test (U = 287.500, z = -0.49, p = .626), nor at post-test (U = 242.50, z = -1.36, p = .174). Data was not normally distributed so non-parametric tests were used, and median scores reported.

Wicoxon Signed Ranks test indicated that neither the pupils in the intervention group (T = 167.50, z = 1.33, p = .183) nor control group (T = 144.00, z = -0.172, p = .864) had statistically significant changes from pre to post-test. Although it did not achieve statistical significance, the increase in the intervention group represented a small to medium effect (r = .19).

Table 3.23 Changes in play same gender peer acceptance for pupils in intervention and control conditions with low initial peer acceptance (PSG)

		N	Median	Range
Intervention (low)	Pre Test	25	0.00	62.12
	Post Test	25	9.09	165.91
Control (low)	Pre Test	25	0.00	91.43
	Post Test	25	-16.67	156.60

3.14.4 Play Different Gender

When considering pupils who started off as low in PDG peer acceptance, the intervention and control groups were not statistically different at pre-test (t (47) = -1.38, p = .173). Data was normally distributed.

There was a statistically significant difference between the gains made by the intervention and control group (t (47) = 4.23, p = .000), with the children classed as low initial peer acceptance in the intervention group making statistically significantly greater gains than those in the control group. This represented a large effect (d = 1.21).

Table 3.24 Changes in play different gender peer acceptance for pupils in intervention and control conditions with low initial peer acceptance (PDG)

		N	*M* at pre test	*SD*	Mean Change Score	SD
Intervention (low)	Pre Test	25	-68.15	15.01		
					36.23	40.31
	Post Test	25	-31.93	40.69		
Control (low)	Pre Test	24	-62.10	15.57		
					-3.73	23.28
	Post Test	24	-65.83	24.18		

3.14.5 A Summary of Findings

Increases in peer acceptance for those children with initially low levels were statistically significantly greater for those in the intervention group than control group in three of the four contexts: WSG (p = .000, d = 1.08), WDG (p = .003, d = 0.92), PDG (p = .000, d = 1.21). These represented large effects. The null hypothesis could be rejected for these contexts.

With regards to PSG, although trends were in the direction of the experimental hypothesis, with the intervention group gaining from pre to post-test, this did not reach statistical significance. The effect size was small to medium (r = .19). The null hypothesis was not rejected.

3.15 A Summary of the Key Findings of the Research (Peer Acceptance and Perceptions of Peer Support)

For ease of reference and clarity, an overall summary of the key findings of the research can be found in Table 3.25.

Table 3.25 A summary of the key findings of the research (peer acceptance and perceptions of peer support)

Research Question	Result	
	Overall	Individual schools
Research Question 1: **Perceptions of Peer Support**	Statistically significantly greater increase in mean change scores for pupils in the intervention condition than control condition *p* = .007 *d* = .42 (small to medium effect)	Trends in all schools were in the direction expected by the experimental hypothesis Statistically significantly greater increase for pupils in the intervention condition than control condition in school C (*p* = .778, d = .10, large effect) No significance in schools A, B or D. Effect sizes ranged from small (*d* = .10 in school B) to large (*d* = .55 in school D)
Research Question 2i.a: Differential impact of CL on pupils' perceptions of peer support Low v medium v high initial peer acceptance	No statistically significant differences between increases in perceptions of peer support for intervention group pupils who had initially low, medium, or high peer acceptance when defined by any of the four contexts (WSG, WDG, PSG, PDG)	

Research Question 2i.b: Differential impact of CL on pupils' perceptions of peer support Low v medium v high initial perceptions of peer support	Statistically significantly greater increases for intervention group pupils with initially low perceptions of peer support when compared to pupils with initially medium or high perceptions of peer support (p = .003, Eta Squared = .13, moderate to large effect)
Research Question 2ii: Differential impact of CL on pupils' perceptions of peer support Low intervention v low control condition (initial perceptions of peer support)	Statistically significant increase from pre to post test for pupils with initially low perceptions of peer support in the intervention condition (p = .001, r = .52, large effect) No statistically significant change from pre to post test for pupils with initially low perceptions of peer support in the control group (p = .131)

Research Question 3: **Peer Acceptance**	Statistically significantly greater increases in mean change scores for pupils in the intervention condition than control condition for: work same gender (p = .000, d = .51, medium to large effect) work different gender (p = .000, d = .74, medium to large effect) play different gender (p = .000, d = .70, large effect) No statistically significant difference found between intervention and control mean change scores for play same gender (p = .398, d = .12, small to medium effect)	WSG individual schools: Statistical significance achieved in school A (p = .014 d = .70, medium to large effect) and school C (p = .002. r = .48, medium to large effect) No statistical significance for school B or D. Small to medium effect sizes (r = .12, d = .45 respectively)
		WDG individual schools Statistical significance achieved in school C (p = .000, d = 2.04, large effect) and school D (p = .016, d = .67, (medium to large effect) No statistical significance for school A or B. Small to medium effect sizes (d = .30, d = .30 respectively)
		PSG individual school Statistical significance achieved in school B (p = .019, r = .33, medium to large effect) and school C (p = .015, r = .38, medium to large effect) However, this was a significant *decrease* for the intervention group in school B No statistical significance for school A or D. Small to medium effect sizes (d = .03, r = .23 respectively)
		(see below)

	When school B was removed from the analysis: Statistically significantly greater increase in play same gender mean change scores for pupils in the intervention condition than control condition (p = .026, d = .37, small to medium effect)	PDG individual schools Statistical significance achieved in school A (p = .020, r = .32, medium to large effect), school C (p = .000, d = 1.96, large effect) and school D (p = .035, d = .58, medium to large effect) No statistical significance for school B (p = .284, d = .31, small to medium effect)
Research Question 4i: Differential impact of CL on peer acceptance Low v medium v high initial peer acceptance	In all four contexts (WSG, WDG, PSG, and PDG), when looking within the same context, pupils with initially lower levels of peer acceptance made statistically significantly greater gains than those with higher initial levels. Often, pupils with medium levels of peer acceptance also made statistically significantly greater gains than those with higher levels, but never greater than those with low peer acceptance. Effect sizes ranged from moderate to large (d = .08) to large (d = .14). *Significant differences found in changes to:* Work same gender peer acceptance scores for intervention group pupils with initially: Low and high WSG peer acceptance (p = .006) Medium and high WSG peer acceptance (p = .010) Work different gender peer acceptance scores for intervention group pupils with initially: Low and high WDG peer acceptance (p = .001) Medium and high WDG peer acceptance (p = .002)	

	Play same gender peer acceptance scores for intervention group pupils with initially: Low and high PSG peer acceptance (*p* = .024) Medium and high PSG peer acceptance (*p* = .003) Play different gender peer acceptance scores for intervention group pupils with initially: Low and high PDG peer acceptance (*p* = .001) Play different gender peer acceptance scores for intervention group pupils with initially: Low and high WDG peer acceptance (*p* = .0016) Medium and high WDG peer acceptance (*p* = .032)
Research Question 4ii: Differential impact of CL on peer acceptance Low intervention v low control condition (initial peer acceptance)	Statistically significantly greater increases in mean change scores for pupils in the intervention condition than control condition for: Work same gender (*p* = .000, *d* = 1.08, large effect) Work different gender (*p* = .003, *d* = 0.92, large effect) Play different gender (*p* = .000, *d* = 1.21, large effect) No statistically significant difference in play same gender scores from pre to post test for pupils in the intervention condition (*p* = .183, *r* = .19, small to medium effect)

4. Discussion

This chapter outlines and critically discusses the findings of the research. Each research question is considered in turn, with links made to relevant literature. Findings are then considered in relation to methodological limitations.

The implications of the research for practice and for future research are discussed. Final conclusions are drawn, summarising the outcomes of the research, and highlighting the unique contribution of the study.

4.1 Findings of the Research: Addressing the Research Questions

The study used a quasi-experimental group design to investigate the impact of CL on pupils' peer acceptance and perceptions of peer support.

Research questions one and two related to perceptions of peer support, as measured using the CSS (Battistich, Solomon, Kim et al., 1995).

Research questions three and four related to peer acceptance, as measured using the SIS (Frederickson & Furnham, 1998).

The findings from each research question will now be considered in turn, with links made to relevant literature.

4.1.1 Research Question One: What is the impact of Cooperative Learning on pupils' perceptions of peer support?

When considering the data from all schools combined, findings suggest that CL is able to have a positive impact on pupils' perceptions of peer support. There were statistically significantly greater gains in pupils' perceptions of peer

support following CL, when compared to pupils in the control group, and this represented a small to medium effect.

There were increases in perceptions of peer support in the intervention group in all individual schools, with effect sizes ranging from small to large. Statistical significance was achieved in school C. Findings at this level indicate that the impact of CL on perceptions of peer support may not occur to the same degree in all contexts. This will be discussed further in section 4.1.9.

4.1.2 Research Question Two: Does the impact of Cooperative Learning on pupils' perceptions of peer support vary, depending on their initial levels of a) peer acceptance or b) perceptions of peer support?

2i- a) Perceptions of peer support and initial levels of *peer acceptance*: There were no significant differences in increases in perceptions of peer support for children with initially low, medium or high peer acceptance in any context (WSG, WDG, PSG, PDG). That is, the degree to which CL impacted upon perceptions of peer support did not appear to be influenced by initial levels of peer acceptance.

2i- b) Perceptions of peer support and initial levels of *perceived peer support:* Pupils who had lower initial levels of perceived peer support made statistically significantly greater gains in perceptions of peer support than their peers with initially medium or high levels. This represented a moderate to large effect.

2ii Low perceptions of peer support (intervention v control): Children with initially low perceptions of peer support had statistically significant increases in perceptions of peer support in the intervention condition from pre to post-test. This represented a large effect. No statistically significant change was found in the control condition.

4.1.3 A Summary of Findings: Perceptions of Peer Support

As described earlier, findings suggest that CL can enhance pupils' perceptions of peer support.

Indications at a whole school and individual school level were consistently in support of the experimental hypothesis. When considering a potential differential impact, the findings of the research were positive in suggesting that CL can be effective in enhancing peer support for pupils with a range of initial peer acceptance levels. Furthermore, for pupils with initially low levels of perceived peer support, CL appeared more beneficial than teaching as usual, and findings pointed to the potential for CL to be effective in 'closing the gap'.

The impact of CL on pupils' perceptions of peer support was small to medium (*d* = .42), and into Hattie's (2009) 'zone of desired effects', with the impact in two schools being medium to large (*d* = .55) and large (*d* = 1.1).

The differences in gains made by pupils with initially low perceptions of peer support, in comparison to their medium and high peers indicated a moderate to large effect size (Eta Squared = .13). Pupils in the intervention group with initially low perceptions of peer support gained significantly more than their control group counterparts, with a large effect size (*r* = .52).

This research builds on a limited evidence base regarding the impact of CL on pupils' perceptions of peer support. Given the lack of literature identified during the systematic literature review, further replication will be important in order to drawn more definitive conclusions.

4.1.4 Links to the Literature: Perceptions of Peer Support

Finding positive impacts on perceptions of peer support following CL is as might be expected from the literature.

Johnson, Johnson, & Anderson (1983) found that, in comparison to their peers in control or comparison conditions, pupils in CL classes had greater perceptions of support from peers. The current research supports this finding, whilst updating and adding to the evidence base.

The finding of statistically significantly greater gains in pupils' perceptions of peer support made by the intervention group when compared to the control group appears in line with Johnson & Johnson's (1999) claim that cooperative relationships are 'a prerequisite to creating a meaningful social support system for each student' (p.64). As suggested by Johnson & Johnson (2013), it is likely that CL facilitated the development of a peer support system, providing opportunities for pupils to be part of a group in which they share resources, support, encourage and praise each other. Although anecdotal in nature, the development of such a peer support system was particularly visible in school C.

Furthermore, literature would suggest that increased social skills following the social skills training, and regular opportunities to practise in real contexts, may have supported pupils in initiating, maintaining and providing social support (Johnson & Johnson, 2013). The degree to which there were real changes in social skills either within CL sessions or beyond was not assessed in the current study. This would be useful to capture in future work in order to triangulate findings and allow for a more detailed understanding of the processes which may be in action.

Additionally, the findings by Johnson, Johnson, Buckman, & Richards (1985) that perceptions of social support increased with increasing frequency and longevity of CL may have some support from the current study. The effect sizes in schools A and B were smaller when compared to schools C and D, with the former schools spending the least time engaging in CL.

Demaray, Malecki, Davidson et al. (2005) wrote that significant changes in levels of perceived social support within an academic year are unlikely. However, the findings of this research indicate that it is possible for changes to occur over a shorter period following CL – although the degree to which this

represents natural, by chance variation or is related to difficulties in measuring perceptions of peer support over a short time scale, may perhaps be questioned. Nevertheless, the findings indicated that pupils at all levels of peer acceptance had gains in perceptions of peer support. Literature suggests that this may have been through a breaking down of the niches of opportunity, providing increased access to social support (Lanier, 2007). Interestingly, this did not appear unique to those with the lowest levels of peer acceptance.

Although the current study did not distinguish between personal and academic peer support, the finding can be viewed as broadly in line with studies by Bertucci, Conte, Johnson & Johnson (2010) and Bertucci, Johnson, Johnson & Conte (2011) included in the systematic review. The CSS (Battistich, Solomon, Kim et al., 1995) appears to contain more questions regarding personal than academic support, suggesting the possibility of an increase in perceptions of personal support as well as academic in the current study. Yet, given the more convincing increases in work peer acceptance when compared to play peer acceptance (see section 4.1.5) in the current study, it is also conceivable that there may have been differences in increases in academic and personal support. As these are both conjecture, further research unpicking any differential effect on the two elements would be useful.

It is useful to acknowledge that increases in perceptions of peer support do not necessarily equate to increases in actual peer support. Further research may look at whether changes in perceptions are linked to changes in actual support. However, as it is the child's felt sense of being supported which is most closely linked to positive psychological outcomes (Wethington & Kessler, 1986), this is perhaps not of primary concern.

4.1.5 Research Question Three: What is the impact of Cooperative Learning on pupils' peer acceptance?

When considering the data from all schools combined, the null hypothesis can be rejected for three of the four contexts. The findings of the current research suggested that CL is able to have a positive impact on pupils' peer acceptance in both work contexts (WSG, WDG), and in one play context (PDG). These represented medium to large and large effects.

At an individual school level, and across contexts, trends were predominantly in favour of the intervention classes, with some significance found in schools A, D and C. At the times when there was no significance found, low power may be considered as an explanation as effect sizes were small to medium. School B appeared to be an exception in the data, with less convincing support for the experimental hypothesis.

In the fourth context, PSG, when all schools were combined, although the intervention group increased in line with the experimental hypothesis, this did not reach statistical significance and the effect size was weak. The null hypothesis cannot be rejected for this context.

In understanding the exception of PSG, it is possible that difficulties with implementation in school B may have contributed to the significant decrease (rather than increase) in the intervention class from pre to post-test on this variable. This is discussed further in section 4.1.9. It may be that, when implemented with higher fidelity, CL is able to have a positive impact on PSG acceptance. Indeed, overall analysis without school B, indicated a statistically significant difference between the changes in scores for intervention and control groups, which represented a small to medium effect. Therefore, the findings that peer acceptance did not increase in PSG should be interpreted in light of data collected and limitations encountered during the research.

As significant changes were found in the PDG context, it does appear that there is possibility for real increases beyond the work context, suggesting again, that in different circumstances, a greater impact and potentially significant finding may have been identified in PSG peer acceptance ratings. However, as there were significant differences at pre-test between intervention and control conditions in PDG, as well as within the intervention grouping, these results should be interpreted cautiously, as should predictions based on them. It may well be that the impacts of CL are most reliably and most likely found within the work context.

4.1.6 Research Question Four: Does the impact of Cooperative Learning on pupils' peer acceptance vary, depending on pupils' initial levels of peer acceptance?

4i In all four contexts, when looking within the same context (WSG, WDG, PSG, PDG), pupils with initially lower levels of peer acceptance made statistically significantly greater gains than those with higher initial levels. Often, pupils with medium levels of peer acceptance also made statistically significantly greater gains than those with higher levels, but never greater than those with initially low peer acceptance. Moderate to large, and large, effect sizes were found.

Although this needs to be taken with caution, as it may be expected that those who start lower in one context may gain more in that same context, it is nonetheless reassuring in suggesting the potential for CL to 'close the gap'. Given literature highlighting the difficulties in intervening with peer acceptance for the least peer accepted pupils, it could be argued that those who start lower may be in that position because of insurmountable difficulties (Bierman, 2004), and therefore would not be expected to have increased as much as their medium, or even high, rated peers.

Interestingly there was some evidence of increases occurring across contexts. Pupils rated as low or medium WDG had statistically significantly greater increases in PDG acceptance than their high WDG peers.

At times, there were relatively large differences between the gains made by low and medium groups, with the low groups gaining most pre to post-test. However, these differences did not achieve statistical significance, which may be a true finding, or may be due to the low power described in section 3.7.

It is also important to note that pupils with high initial peer acceptance levels did, at times, appear to decrease from pre to post-test. This may have been due to factors such as inflated pre-test scores due to overly positive reputational biases, regression to the mean, or ceiling effects.

4ii Furthermore, increases in peer acceptance for those children with initially low levels were statistically significantly greater for those in the intervention group than control group in both work contexts (WSG, WDG), and in one play context (PDG). These represented large effects.

With regards to PSG, although trends were in the direction of the experimental hypothesis, with the intervention group gaining from pre to post-test, this did not reach statistical significance. The effect size was small to medium. This is perhaps as reflected in the overall findings that CL did not appear to have a significant impact on PSG peer acceptance. It may also be due to the low power described in section 3.7, with up to a 59% chance of failing to detect a medium effect, and up to an 89% chance of failing to detect a genuine but small effect.

4.1.7 A Summary of Findings: Peer Acceptance

The current research suggests that CL can significantly enhance pupils' peer acceptance, particularly in the work context. Findings for the play context were

mixed, with statistically significant increases found in PDG but not PSG peer acceptance. Possible reasons for this finding have been discussed.

The impact of CL on pupils' peer acceptance in the WSG, WDG, and PDG contexts indicated medium to large effects (*d* = .51, *d* = .74, *d* = .70 respectively), and small to medium for PDG (*d* = .12). In three contexts, effect sizes were well into Hattie's (2009) 'zone of desired effects'.

When considering a potential differential impact, the findings suggest that the impact of CL may vary, according to pupils' initial levels of peer acceptance. This was in favour of those pupils with initially low or medium levels of peer acceptance across all contexts. The differences in gains made by pupils with initially low or medium perceptions of peer acceptance, in comparison to their high peers indicated moderate to large or large effect sizes.

Furthermore, the least accepted pupils gained significantly more in intervention than control classes in both work contexts (WSG, WDG), and PDG peer acceptance. These represented large effect sizes for WSG, WDG, and PDG contexts (*d* = 1.08, *d* = .92, *d* = 1.21 respectively), and a small tg medium effect size for PDG (*r* = .19).

4.1.8 Links to the Literature: Peer Acceptance

The findings of the research are supportive of previous literature suggesting that CL holds the potential to enhance pupils' peer acceptance.

Findings are in line with studies such as Piercy, Wilton, & Townsend (2002) and Putnam et al. (1996) which have reported increases in peer acceptance when focusing on mainstream peers rating special population peers. However, whereas several studies in the systematic literature review found no significant changes in peer acceptance for mainstream pupils rating their mainstream peers from pre to post-test (e.g. Andre et al., 2011; Andre et al., 2013; Topping

et al., 2011), the current study suggests that there *can* be an impact in mainstream populations.

As with perceptions of peer support, the findings of the current study are particularly encouraging given the general belief that peer acceptance is relatively stable (Zettergren, 2005). The findings are optimistic in suggesting potential for change for children with low peer acceptance, particularly in work contexts, following CL.

As findings were not found consistently across all four contexts or across all four schools, the research highlights a need for caution in the application of CL, as well as pointing to further work in understanding these inconsistencies and ways in which to support or manage them. The inconsistencies are in fact in line with other literature which has found that any positive findings at an individual class level may be almost equally balanced out by negative findings in other classes (e.g. Jacobs et al., 1996; Topping et al., 2011).

The 11 week intervention period was a pragmatic choice, linked to term length and research timescales. Due to further school specific demands, this was not achieved in any class. It is therefore possible that inconsistent findings, particularly for PSG, and for school B, could be due to the short duration of the intervention. Indeed, Gillies & Ashman (1997) argue that it is unreasonable to expect a change in norms or behaviours for either teachers or pupils in a short space of time. Rather than indicating that an intervention has not been successful and that there has not been an impact on pupils, a lack of positive outcomes in quantitative measures may indicate that the intervention was of too short duration for this to have been measured (Cohen, Manion, & Morrison (2011). It may be that for convincing generalisation to the play context, more time is needed.

The current study found most support for impact in the work context, which may appear logical, as the CL groups were within a work rather than play context. This is not a unique finding, as significant increases in work peer acceptance were reported by Putnam et al. (1996) following CL. However, literature is often

inconsistent regarding the contexts in which CL may have most influence on peer acceptance. Craig (2010) found effects across both work and play contexts, with same gender acceptance scores showing most impact. On the other hand, Topping et al (2011) reported differences only in play contexts – although the use of post-test only data at this point indicates caution in the interpretation of this finding.

With regards to peer acceptance, the current study can be viewed as a partial replication of that of Craig (2010). With Coolican (2009) noting that the chance of gaining positive results in more than one study in the absence of a genuine phenomenon is highly unlikely, the study provides further support for the impact of CL on peer acceptance.

In doing so, the current study also builds upon an emerging evidence base regarding the Kagan model of CL, going some way to addressing Johnson & Johnson's (2002) caution regarding a lack of evidence for many models of CL.

4.1.9 Additional Findings: Differences Between Schools

Due to the potential for multiplicity, the individual school results have not been given undue attention. However, given the assertion of Wright (2006) that statistics is about communicating patterns in the data, rather than restricting discussion solely to specific hypotheses, two interesting patterns which emerged in schools B and C are worthy of further discussion.

These relate primarily to issues around intervention fidelity. Low implementation fidelity 'potentially weakens outcomes, makes interventions appear ineffective, [and] masks mechanisms of effect for interventions' (Butel et al., 2015, p.461). The degree of implementation fidelity achieved may explain the differences in findings between schools – and may have impacted on overall findings and conclusions.

Although intervention adherence in school B was good, the exposure was lower than in the other three schools, with CL appearing to be the exception rather than the norm. On the whole, the findings from school B were non-significant and often showed weak, or small to medium effects, as might be expected given the low exposure. Additionally, there was a decrease in both WDG and PSG scores in the intervention group pre to post-test, with this reaching statistical significance for PSG.

It is useful to consider the possibility borne out in this study that a limited application of CL may have less impact – and, at times, may have a negative impact. Novelty and disruption effects, whereby negative results occur due to initial intervention disruption, may have impacted on findings. Furthermore, a change of teacher during the intervention may have impacted on results in school B, and leads to caution in interpreting results for this school.

This may be have been compounded by inconsistencies in the presence of positive interdependence, small group skills and group processing as highlighted during intervention adherence checks. Without these aspects, Battistich, Solomon, & Delucchi (1993) warn that there is an increased risk of negative interaction which could lead to decreases in peer acceptance and peer liking.

Furthermore, anecdotal data captured during intervention fidelity checks indicated that the teacher in school B appeared to retain a relatively high degree of control over pupil-pupil interactions, and closely managed each CL activity, admonishing children who were not achieving as she may have wished during their cooperative efforts. With studies suggesting that teachers can have their own impacts on peer ecology and peer acceptance through the messages they communicate to pupils and their peers, it is possible that the impact of CL on peer acceptance was lower in this intervention class as teacher behaviour may have inadvertently reinforced the differential status of pupils (Harrist & Bradley, 2003). Without perceived equality of status, positive impacts of CL on peer acceptance may not be likely (Jules, 1991).

Similarly, with relatively low time to interact freely and relatively low amounts of pupil-pupil dialogue, it is possible that pupils in school B did not have sufficient time or opportunity to build upon existing peer relationships (Gillies, 2015).

Despite some potential difficulties, the decision was taken to include school B in the overall analyses as such a variation is likely to be representative of the somewhat compromised application of CL in many real world educational contexts.

On the other hand, intervention adherence in school C was good, continued to improve throughout the study, and exposure was high. CL was used at least twice daily, characterised by short bursts, and often used multiple times within a lesson. CL became the norm in this class. In school C, pupils remained with their class teacher all day, whereas in school B, pupils moved into different classes for maths. It may also be relevant to note that the class teacher in school C was an NQT, whereas the teacher in school B had taught for 20 years.

In contrast to school B, school C had statistically significant differences on each DV measure, with large or medium to large effects. This may be linked to the delivery model, or may in part be explained by some relatively low pre-test scores in the school C intervention class – although relatively low pre-test scores were also found in the school C control class, rather than solely in the intervention group.

It does appear that, on balance, in a school where perhaps there is relatively low peer acceptance and low perceptions of peer support, CL may be particularly effective in increasing both aspects – and, in this instance, with the given delivery model, appeared vastly superior to the ‘teaching as usual’ wait list control condition.

The differences between schools, but most notably between school B and C, needs to be taken in to consideration when interpreting overall findings. The inclusion of school B in overall analyses may mean that the positive effects of CL across the remaining schools where intervention delivery was more close to

ideal, has been underestimated. Alternatively, the influence of school C may have led to the overall effect of CL being over exaggerated.

The differences between schools adds a level of caution to the findings, as although findings of the study were predominantly positive, success on an individual class level cannot be guaranteed. However, it is perhaps unreasonable to expect even the most efficacious intervention to work to the same extent in all contexts, particularly given a potentially wide range of extraneous variables. This does not mean that an intervention is not useful, and not having an impact, but does point to a need to further understand these variables, and identify why, and for whom CL might be most effective. This may involve a closer study of differences in teacher characteristics and classroom environment, as MacMillan & Morrison (1984) highlight aspects such as teacher reinforcement patterns and teacher attitudes towards individual pupils in influencing social acceptance patterns within a classroom.

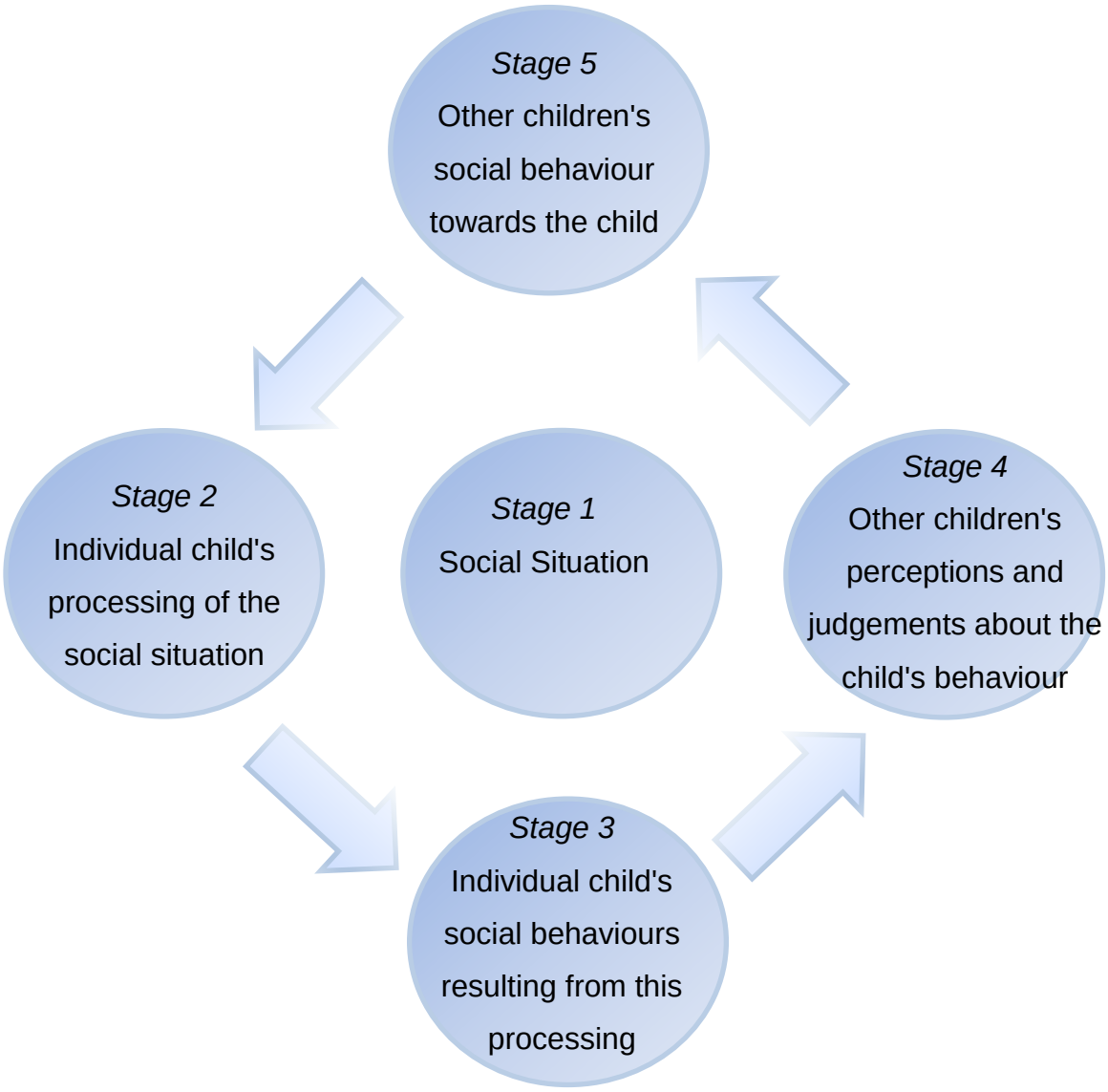
4.2 Links to the Literature: Potential Mechanisms of Impact

When discussing the four research questions, for both peer acceptance and perceptions of peer support, literature has been presented in order to set findings into a wider context. More general theoretical and procedural aspects will now be discussed.

4.2.1 Theoretical Aspects

The Dodge et al. (1986) model of social functioning and social exchange was highlighted in section 1.4 as able to describe the cyclical and complex processes which may contribute to social acceptance or rejection, and access to peer support.

Figure 4.1 A model of social exchange in children (adapted from Dodge et al., 1986)



The model suggests that CL may have impacted upon:

- a child's processing of a social situation, leading to change in their subsequent behaviour.
- changes in other children's perceptions and judgements about a child, leading to changes in their behaviour towards that child

An example of how CL may have supported increases in peer acceptance and perceptions of peer support is described below. It is important to note that this is suggestive rather than definitive.

Having been involved in CL over a period of weeks, a child's perception and understanding of a social situation at stage 2 may begin to change. Having engaged in supportive and positive experiences promoted by CL structures, the child may become more accurate in their processing of social cues (Price & Dodge, 1989), more likely to expect positive interactions, and more likely to interpret the behaviour of others as supportive (Lakey & Cohen, 2000).

As a result, the child's behaviour at stage 3 may move away from Bierman's (2004) social toxin behaviours, with the child displaying increasingly prosocial, supportive and attentive behaviours. Direct social skills training (Gillies, 2016), and opportunities to learn from their more socially adept peers (Kagan & Kagan, 2009) may support an increasing range and quality of positive behaviours.

CL seeks to ensure everyone can make a valued contribution, with differences in competence minimised, and, through individual accountability and positive interdependence, limits opportunities for hitchhiking or social loafing. As such, and with improved social behaviours at stage 3, the peer group may begin to perceive the child in a more positive light, with frequent and repeated positive interactions during CL challenging negative reputational biases at stage 4 (Frederickson & Cline, 2009).

As these biases are broken down, the peer group may respond in increasingly prosocial ways to the child (stage 5), which supports further ongoing change for the child at stage 2, as now that they have experienced success in meeting social goals, their patterns of thinking and behaviour become more likely to support their inclusion and acceptance (Toner & Munro, 1996). As the child's access to the peer group increases, they have increased opportunities to engage in high quality, positive interactions, with further opportunities to enhance their social skills, and to seek support (Lanier, 2007).

Given the circularity inherent in the model, it is likely that impacts are occurring at all levels simultaneously, with changes at any one stage magnifying and increasing the likelihood of changes at any other stage. Whilst the example given above begins with the individual child, the description could have started at any of the stages.

The Dodge et al. (1986) model perhaps also helps in understanding why impact was more consistently found within the work rather than play context in the current research, given the different nature of the social situation (stage 1). It is possible that improved prosocial behaviours were restricted to the clear structures and expectations of CL group work. Similarly, it is possible that any changes in reputational biases were greater in the work context, as that was where the challenges to reputations were being made, rather than in the playground. It would be useful for future research to investigate this further, for example, looking at whether any changes in social behaviours observed during CL sessions are generalised to other areas of the taught curriculum, as well as to other contexts, such as the playground.

As the evidence base builds for the Kagan model of CL in enhancing peer acceptance and perceptions of peer support, it would be useful for further research to investigate changes at each level. That is, to what extent and in what ways have the social behaviours of an individual child and the peer group changed?

The work of Lanier (2007) suggests that pupils who are isolated from their peers have limited avenues for seeking social support. CL, by its very nature, brings immediate changes in the amount of time children spend truly working together, and, perhaps, for less accepted children who may have experienced limited social opportunity (Bierman, 2004), immediately increases their access to peers. The current study found that whilst all intervention classes made gains in perceptions of peer support, classes did not all make gains in all of the four peer acceptance contexts. It is possible that perceptions of peer support are more susceptible to change, and at an earlier stage, than is peer acceptance, which may require further influences and interactions over a longer period of

time. The development of a peer support system through face to face promotive interaction facilitated through CL lends some support to this notion (Johnson & Johnson, 2013).

For pupils in intervention conditions, literature would suggest that as their inclusion is facilitated through increases in peer acceptance and perceptions of peer support, so to would be their feelings of belonging (Baumeister & Leary, 1995). With Maslow (1970) viewing a sense of belonging as an important stepping stone towards further wellbeing, it is conceivable that impacts in the current study extended beyond the DVs measured.

4.2.2 Procedural Aspects

As noted by Johnson & Johnson (2002), CL is essentially a generic term, and the actual detail of any intervention approach may vary greatly. It is therefore useful to reflect on elements of the intervention in the current study which may have relevance to the findings reported.

The Kagan approach was selected as literature indicated that a relatively direct approach is easier for teachers to pick up, requires less training time, and is more easily implemented (Johnson & Johnson, 2002). Whilst essential principles and processes were usually clear within CL sessions, even with the Kagan approach, there was some evidence of teachers modifying approaches in ways which detracted from the essential principles, and therefore may have reduced impact (Antil et al., 1998).

Literature highlighted the importance of social skills and group processing (e.g. Gillies, 2016). All intervention classes had a distinct period of social skills training prior to CL, which literature suggests would support increases in pupils' peer acceptance through increasing prosocial behaviours, as well as supporting children in both initiating and accepting peer support (East, Hess, & Lerner, 1987). Group processing, however, although usually present, was not consistently so, even within a class. Future training should be adapted to

ensure that the benefits of and approaches to supporting group processing are given greater emphasis as it is perhaps this area which is most easily overlooked by teachers.

Additionally, the current study aimed to investigate the impact of CL on pupils' peer acceptance and perceptions of peer support given training at a level which could be readily and realistically provided by an EP, rather than at a level which would be most likely to secure positive findings from a research perspective. Contrary to studies which have relied upon extensive training and support packages over a period of time, this study does indicate that relatively brief training in Kagan structures can promote change, and that this can occur with relatively short intervention duration. Further work to understand the timescales regarding such changes would help to refine an intervention model.

Finally, Petticrew & Roberts (2006) note that understanding the extent to which changes are maintained over time is critical in understanding the effectiveness of an intervention. However, the extent to which changes in the current study may be maintained is unclear. Due to the timing of the current study (Summer term), it was not possible to conduct a follow up with intact classes or with the same teachers. Whilst Jacques, Wilton, & Townsend (1998) found enhanced social acceptance scores of focus pupils to be maintained at a five week follow up, this was the only study identified in the systematic literature review which included a follow up. With a study by Frederickson, Warren, & Turner (2005) finding that peer acceptance scores following a 'circle of friends' intervention returned to baseline in the term following the intervention, it will be important for future research to investigate the extent to which impacts of CL are maintained.

4.3 Methodological Limitations

As discussed in section 2.1.2, the post positivist approach and quasi-experimental design methodology allowed for research in uncontrolled, real world settings (Mertens, 2015). However, such real world research brings with it a range of challenges, with findings always imperfect and fallible (Robson,

2011). Issues of internal and external validity, and reliability identified as particularly pertinent to the current study will therefore now be discussed.

4.3.1 Internal Validity

Selection: Levels of consent

In accordance with the University of Nottingham ethical guidance, parental consent for children to complete the measures was gained through an opt in process. As highlighted in section 2.5.3, this meant that, at times, rates of consent were low.

Whilst this is perhaps not unusual, with Gifford-Smith & Brownell (2003) noting that participant and consent rates are usually around the 70-75% range in studies which use sociometric measures, it does raise concern that low rates of consent may have led to an under or over estimation of the impact of CL, particularly if there was something salient about the pupils who did not complete the measures. Indeed, Noll et al. (1997) suggested that non-participation in sociometric data collection may be systematic, with children without parental consent described by their peers as being less accepted, more aggressive, and less academically able. Although teachers in the current research did not note any similarities or differences between pupils who did and did not consent to completing the measures, there remains potential for an unexpected and unidentified group similarity to be present.

However, the extent to which this may have been problematic is unclear, as McKown, Gumbiner, & Johnson (2011) argue that high levels of non-participation can be acceptable. With participation rates as low as 40%, they found social preference scores to be highly correlated (above .80) with those when complete data was used.

Selection: Existing differences at pre-test

There were some existing differences at pre-test, both between classes within a school, and within the overall PDG grouping. Unfortunately, as noted by Miller & Chapman (2001), 'there is no statistical method that can address the question of whether two groups that differ on variable A would differ on variable B if they did not differ on variable A' (p.43). Instead, all that can be done is to interpret findings in light of these pre-existing differences (Miller & Chapman, 2001).

Lower scores at pre-test may mean a greater likelihood of increases in scores, regardless of intervention effectiveness, with impact overestimated for these classes. On the other hand, it could be argued that those classes and pupils who were lower at pre-test were lower due to a genuine, long standing and resolute deficit, which would be harder for an intervention to impact upon, meaning any impact may be under estimated.

As noted by Rosenbaum (1987), difficulties associated with differences at pre-test can be best accounted for through the aggregation of a number of individually compromised studies, none of which may have random assignment but collectively present a strong inferential case. As such, despite some internal validity concerns, the findings of the current study are nonetheless important.

Regression to the mean

Regression to the mean is of particular concern when investigating the differential impact of CL, as issues are compounded by categorising pupils into groups based on their baseline scores. Increases for the low peer acceptance group in comparison to the high peer acceptance group therefore need to be interpreted with some caution.

Similarly, when considering small increases in scores for high peer accepted pupils, as well as some indications of decreases in scores, regression to the mean should be considered as an explanation, as should the impact of potential ceiling effects.

Intervention fidelity and adherence

Please refer to section 4.1.9 for discussion about the ways in which intervention fidelity and adherence may have impacted on the findings of the study.

Additionally, although work by Webb & Cullican (1983) suggest that there is no impact on group interaction between groups which stay together and groups which change, it would nonetheless have been useful for intervention monitoring to have included information on how often different teachers changed the groupings in their classroom. Similarly, although heterogeneous groupings were discussed in training, these were not tracked in detail during the intervention period. Information on both of these aspects could have been useful in understanding school level differences further.

4.3.2 External Validity

The impact of CL was not the same across all intervention classes, and this raises questions around the external validity of findings. That is, different teachers, in different schools, with different classes, gained different outcomes and therefore there may be some difficulties in attributing impacts of the CL solely to the intervention itself or in generalising it to other contexts.

4.3.3 Reliability

Potential difficulties associated with the approaches taken to measuring the DVs will now be considered.

It is possible that pupils' responses may have fluctuated from pre to post-test due to influences other than the intervention, such as emotional affect, motivation or social desirability bias. Due to the age of pupils and focus of the DVs, it is possible that those in the intervention classes may have responded positively at pre-test, and, having had opportunities to truly work with their classmates, responded more realistically at post-test. Overly positive scores at

pre-test may therefore have masked an effect. Nevertheless, information gained from peers in this manner provides unique and valuable data which cannot be gained from other sources (Gresham & Stuart, 1992).

There are a number of different ways of measuring peer acceptance, with the most common methods using peer nominations or peer ratings. Whilst Gifford-Smith & Brownell (2003) argue that the findings from both methods are 'reasonably comparable' (p.240), this view is not shared by all, with Frederickson & Furnham (1998b) arguing that the method chosen may lead to differing findings. Therefore, whilst there is undoubtedly a role for peer nominations and peer ratings, rather than relying exclusively on such measures to evaluate impact, future work may consider integrating data from multiple sources and informants in order to be more confident in the reliability of findings (Renk & Phares, 2004). However, difficulties with the utility and validity of teacher reports of peer acceptance should always be considered (Parker and Asher, 1987).

Furthermore, although literature suggested that both peer acceptance and perceptions of peer support would be stable over the course of the study, there were changes in some of the control groups, as well as in the intervention groups. This highlights concerns regarding the reliability of measuring these constructs over a short period of time (Murphy, Grey, & Honan, 2005).

A further difficultly arose from the use of the CSS itself as a measure. Some children appeared to find the reversed and negative questions difficult to interpret, leading to some uncertainty as to whether they were able to accurately and reliably select their intended option. Clarification of the questions was given, and pupils were encouraged to ask for further support if needed. However, in a whole class situation, it is unclear how effective these steps were, and any confusion may have impacted on the findings. In the future, small group administration of the CSS would be useful in order to ensure reliability of responses.

4.3.4 Limitations in Understanding the Differential Impact

A number of approaches to analysing the differential impact were considered, with a group design selected as most appropriate in addressing the research question at this point.

However, although the interquartile range was used as a recognised statistical approach which allowed impartial groupings, with such a group design, there is always potential for findings to be an artefact of the grouping method. It may be that, in the future, more advanced regression methods are considered to supplement understanding.

Additionally, there may have been a confounding of results as low pupils were identified across the whole sample, rather than in individual class groups. This meant that the number of low pupils contributed by each school varied, and therefore findings may have been unduly influenced by the level of impact in any given class. Furthermore, low power may have meant that significant differences were not identified.

With previous research noting that peer acceptance levels of the least accepted pupils are particularly stable (Jules, 1991), or that these pupils may be increasingly rejected following CL (Garel, 2001), the findings of the current study are interesting in finding positive changes for this group of pupils. Whilst this may reflect true findings, in line with the post positivist paradigm, it is important to consider other explanations for the results.

Literature indicates that 10% to 16% of pupils may have difficulty in gaining peer acceptance (Bierman, 2004; Terry & Coie, 1991). It is possible that using a relatively large percentage of children in the low peer accepted group in the current study may have led to the group being overly heterogenous, masking patterns for the most rejected children. However, there was a need to balance this with the potential to overlook the impact for children who were less rejected but nevertheless at risk for negative outcomes (Terry & Coie, 1991). Whilst raw

data does not indicate it to be the case, it remains possible that some impact (or lack thereof) for the most rejected was hidden within the wider low peer accepted grouping. Indeed, when looking at the data in the differential analyses (4ii), the standard deviations and ranges of data for low peer accepted pupils in the intervention condition were sometimes noticeably greater at post-test than at pre-test – although this was also, at times, found in the control condition.

When comparing gains made by pupils with initially low perceptions of peer support and peer acceptance to their medium and high peers, potential impacts of regression to the mean, and ceiling effects should be considered when interpreting impacts. However, analyses comparing pupils with initially low perceptions of peer support and peer acceptance in the intervention and in the control condition goes some way to addressing such concerns (Schagen & Hodgen, 2009).

Whilst imperfect, the findings of the current research are nonetheless interesting, and provide a platform for future work using different methodologies in order to gain a more detailed picture of the impact of CL on the least accepted pupils. Suggestions are given in section 4.4.2.

4.4 Implications of the Findings

In considering the impact of CL on pupils' peer acceptance and perceptions of support, the current study raises implications for the use of CL in practice, and serves to generate further areas which would benefit from investigation. These will now be discussed.

4.4.1 Implications for Practice

The research has been effective in contributing to the evidence base regarding the impact of CL on pupils' peer acceptance and perceptions of peer support, and has reported predominantly positive findings. Furthermore, as the study did

not find a differential impact which disadvantaged the least accepted pupils or those with the lowest perceptions of peer support, those recommending CL to schools as an approach to support such pupils can apparently do so with some confidence.

Rather than viewing peer acceptance and perceptions of peer support as stable over time and resistant to change, CL appears to, at times and in certain contexts, be able to promote opportunities for change. This should provide hope not only to children themselves, but also to school staff and other professionals. Indeed, if CL was incorporated in to the EP toolkit, it may be an appropriate intervention and support for children whose presenting difficulties may be viewed as arising from the long term detrimental impact of poor peer acceptance and lack of peer support (Baumeister & Leary, 1995).

When considering implementation, the study indicates that the use of Kagan structures as an introduction to the principles and practice of CL is readily accessible for class teachers, without unsustainable or unrealistic amounts of training or ongoing support. Similarly, although no classroom could be said to have implemented CL at a perfect level, positive effects were nonetheless found. This should be reassuring and encouraging to teachers considering using CL in their practice. Indeed, although not explored directly in this study, feedback proffered by teachers in the current research was overwhelmingly positive – with all schools planning to develop CL approaches further. Future work may therefore look at the development and impact of whole school approaches to CL.

The research has also indicated that frequent, short bursts of CL activity can be effective, and that changes are possible over a relatively short intervention period. This is useful information to support a delivery model, although further empirical investigation would be beneficial.

Nevertheless, differences in the impact of CL between classes does highlight a continued need to understand the specific contexts in which CL can be *most* effective. Similarly, although a short amount of training appeared sufficient for

most teachers, it is not necessarily enough for all. Ensuring opportunities for follow up support is likely to be important in refining classroom practice and increasing the likelihood of positive impacts.

As well as a remedial intervention, there is potential for CL to be used as a pedagogical approach and preventative strategy. Thus, there is a role for the EP in developing and delivering training, and supporting practice, not only in individual casework and at a school based level, but also at the initial teacher training level. Indeed, the study suggests potential for positive outcomes for a wide range of pupils, not just those who may have been referred to the EP as part of a deficit based casework approach or with initially low levels of peer acceptance. It may also be that internalising CL as a pedagogical approach rather than as an 'add on intervention' would be easier for teachers new to teaching, rather than those who have established their own practices over many years.

Finally, the research reminds one of the importance of taking an ecological approach, intervening at both individual and peer group levels, in order to enhance peer acceptance and perceptions of peer support.

4.4.2 Implications for Future Research

Whilst it is clear that this research has added to the current evidence base, it has also thrown up further questions which would be fruitful to explore. These have been highlighted throughout the chapter in an ongoing manner, and have included questions around:

- What are the general teacher characteristics and practices which influence the impact of CL on peer acceptance and perceptions of peer support?

- What patterns of delivery are most effective in increasing peer acceptance and perceptions of peer support?

- To what extent and in what ways have the social behaviours of an individual child and the peer group changed? Is there a relationship between increases in social skills and increases in peer acceptance and perceptions of peer support?

- What is the optimal intervention length? Have main effects been established by week 5, week 8, week 11? Will improvements continue to be seen?

- Does the impact of CL last beyond the active intervention period? At what rate does it diminish? How can impacts be maintained?

In order to overcome concerns regarding findings being an artefact of the measure selected, triangulation of data within a study would be useful. This may be through alternative or additional sociometric measures (e.g. sociograms) or observations (e.g. of specific cooperative behaviours).

Within this, further research which focuses on developing the evidence base for Kagan models would be useful.

Further avenues for potential research are summarised in Table 4.1.

Table 4.1 Suggestions for future research

Suggestions for a Research Question	Rationale	Possible Design
What are the individual experiences of children and staff of CL? What are their perceptions as to how CL may have influenced peer acceptance and perceptions of peer support? Does this vary according to pupils' initial levels of peer acceptance?	Griffiths, Stenner, & Hicks (2014) highlight the importance of consulting with children when evaluating the effectiveness of interventions designed for their own wellbeing. This is an area overlooked in the current study. Taking a constructivist approach to investigate and illuminate the lived reality of the intervention for pupils - and staff - through detailed descriptions of their experiences would be a useful addition. Such an approach may support in illuminating more closely the ways in which CL impacts upon different aspects of the Dodge et al. (1986) model.	Case studies Narrative approaches, interviews and focus groups Mixed methods
What is the impact of CL on the peer acceptance and perceptions of peer support of those few children who are extremely rejected by their peers?	The current study suggests generally positive impacts of CL on children with low peer acceptance and perceptions of peer support. However, an individualised look at the impact of CL on extremely low peer accepted children would be helpful.	SCED Case study

What is the impact of CL on the perceptions of peer support of specific special population pupils?	Studies indicate that certain groups may perceive less support than their peers, e.g. those with 'ADHD type behaviours' (Demaray & Elliott, 2001) or learning difficulties (Wenz-Gross & Siperstein, 1998). Understanding whether CL can support these pupils effectively would be useful. Given the often transient nature of class groups in Pupil Referral Units, understanding the potential of CL in those contexts could also be useful.	SCED Case study
What is the impact of CL on pupils' perceptions of teacher support?	Social support can be gained from multiple sources, including teachers (Demaray, Malecki, Davidson et al., 2005). As pupils in the study by Johnson, Johnson, & Anderson (1983) noted greater perceptions of support from teachers, that would perhaps be an additional useful measure or angle to take in the future.	Quasi-experimental
What is the relationship between peer acceptance and perceptions of peer support? Do perceptions of peer support increase as peer acceptance increases?	Further work investigating the relationship between the two constructs of peer acceptance and perceptions of peer support would be useful. Although the Dodge et al. (1986) model and literature presented in chapter 1 suggests the likelihood of a relationship, few studies have explicitly considered this link. East, Hess, & Lerner (1987) was one such study, and this indicated it to be an area fruitful for further work.	Quasi-experimental Correlational

Are there other significant variables within this relationship?		
How does CL impact on social networks within a classroom? What is the impact on social network 'isolates'?	Social cognitive mapping techniques could be used to assess the impact of CL on social networks, an overlapping yet distinct dimension of the peer group experience (Gifford-Smith & Brownell, 2003). As the current study had difficulties with levels of opt in consent, this approach could be advantageous. The social cognitive approach represents the consensus of members of the peer group, meaning that it is perhaps less critical to obtain data from every member of the class.	Quasi-experimental

4.5 Summary and Conclusions

The current study used a quasi-experimental group design to investigate the impact of CL on pupils' peer acceptance and perceptions of peer support, two primary indicators of social inclusion in a school context.

The study indicated that CL can enhance pupils' perceptions of peer support, and that the degree to which it is effective is not influenced by initial levels of peer acceptance. Pupils with initially lower levels of perceived peer support made greater gains in perceptions of peer support than their classmates, and these gains were also greater than those made by pupils in control classes.

The study indicated that CL can enhance pupils' peer acceptance, particularly in the work context. Findings for the play context were mixed, with statistically significant increases found in different gender but not same gender peer acceptance ratings. Possible explanations for this have been discussed. There appeared to be a differential impact of the intervention, in favour of pupils with initially low or medium levels of peer acceptance. The least accepted pupils gained significantly more in the intervention condition than in the control condition in both work contexts and in the play different gender peer acceptance context.

Although some caution is advised in interpreting results – particularly since the degree of impact was not consistent across individual classes – it is concluded that Cooperative Learning approaches can have a valuable impact on pupils' peer acceptance and perceptions of peer support, with the potential ability to 'close the gap' on these constructs.

When considering effect sizes, and seeking to draw a distinction between statistical and practical significance (Dancey & Reidy, 2014), the findings of the current study are positive, with the majority of findings overall - as well as impacts in individual schools - exceeding Hattie's (2009) 'zone of desired effects', with many medium to large or large effects reported. Given that the

current study sought to typify rather than optimise the delivery of CL, these findings are encouraging in indicating the potential for CL to impact on pupils' perceptions of peer support and peer acceptance.

The exception to this, once more, appeared to be school B, where effect sizes were generally smaller. However, this should be understood alongside difficulties with implementation fidelity in school B, and, given literature indicating that intervening with perceptions of peer support and peer acceptance is difficult, should also be considered in light of the writings of Higgins, Kokotsaki, & Coe (2012) who note that even 'small' effects may be considered educationally important and of practical significant when impacts arise from easy or cheap interventions, or with groups where outcomes are hard to influence - such as pupils struggling with gaining peer acceptance and peer support.

Within the post-positivist paradigm, it is acknowledged that findings are 'true' only within certain parameters of probability, and within the conditions which were encountered during the study (Mertens, 2015). As such, limitations in the current study have been discussed, and implications for practice and future research provided.

4.5.1 Reflections on the Unique Contribution of the Study

The primary unique contribution of the current research is in looking at the impact of CL on perceptions of peer support, building on limited and dated research in this area.

The research also makes a clear contribution to the existing evidence base for the impact of CL on peer acceptance and perceptions of peer support by:

- the use of a Kagan model
- the use of a UK context
- the use of a teacher led, intact mainstream class delivery model
- exploring the differential impact

In so doing, the study has provided findings which are relevant to EP practice, particularly in highlighting the possibilities for CL in supporting peer acceptance and perceptions of peer support in a low cost, minimal disruption, subtle and inclusive manner. Furthermore, it has addressed concerns that such an approach may widen the gap between the least accepted children and their peers, and result in a 'Matthew effect' of the rich getting richer. There was no evidence of this occurring.

It has also suggested that some caution is necessary in the implementation of CL, as impact may be different in different contexts, and may not be consistent at an individual class level. Consideration has been given to these aspects.

5. References

Abrams, D., Hogg, M., & Marques, J. (2005). A social psychological framework for understanding social inclusion and exclusion. In D. Abrams, M. Hogg, & J. Marques, *Social Psychology of Inclusion and Exclusion* (pp. 1-24). New York: Psychology Press.

Alexander, R. (2008). *Towards Dialogic Teaching: Rethinking Classroom Talk.* Cambridge, UK: Dialogos.

Allison, P. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology, 20*, 93-114.

Allison, P., & Pomeroy, E. (2000). How shall we "know?" Epistemological concerns in research in experiential education. *The Journal of Experiential Education, 23*(2), 91-98.

Anan, R., & Barnett, D. (1999). Perceived social support mediates between prior attachment and subsequent adjustment: A study of urban African American children. *Developmental Psychology, 35*(5), 1210-1222.

Andre, A., Deneuve, P., & Louvet, B. (2011). Cooperative learning in physical education and acceptance of students with learning disabilities. *Journal of Applied Sport Psychology, 23*, 474-485.

Andre, A., Louvet, B., & Deneuve, P. (2013). Cooperative group, risk taking and inclusion of pupils with learning disabilities in physical education. *British Educational Research Journal, 39*(4), 677-693.

Antil, L., Jenkins, J., Wayne, S., & Vadasy, P. (1998). Cooperative learning: prevalence, conceptualisations, and the relation between research and practice. *American Educational Research Journal, 35*(3), 419-454.

Armstrong, R. (2014). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics, 34*, 502-508.

Baines, E., Blatchford, P., & Chowne, A. (2007). Improving the effectiveness of collaborative group work in primary schools: Effects on science attainment. *British Educational Research Journal, 33*(5), 663-680.

Baldwin, W. (1999). Information noone else knows: The value of self-report. In A. Stone, C. Bachrach, J. Jobe, H. Kurtzman, & V. Cain, *The Science of Self Report* (pp. 3-8). New York: Psychology Press.

Bandura, A. (1971). *Social Learning Theory.* New York: General Learning Press.

Barnett, A., Van der Pols, J., & Dobson, A. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology, 34*(1), 215-220.

Barrett, W., & Randall, L. (2001). Investigating the circle of friends approach: Adaptations and implications for practice. *Educational Psychology in Practice*, 353-368.

Battistich, V., & Watson, M. (2005). Fostering social development in preschool and the early elementary grades through cooperative classroom activities. In C. Gillies, & A. Ashman, *Cooperative Learning: The Social and Intellectual Outcomes of Learning in Groups* (pp. 19-35). London: Routledge Falmer.

Battistich, V., Schaps, E., Watson, M., Solomon, D., & Lewis, C. (2000). Effects of the Child Development Project on students' drug use and other problem behaviors. *Journal of Primary Prevention, 21*, 75-99.

Battistich, V., Solomon, D., & Delucchi, K. (1993). Interaction processes and student outcomes in cooperative learning groups. *The Elementary School Journal, 94*(1), 19-32.

Battistich, V., Solomon, D., Kim, D., Watson, M., & Schaps, E. (1995). Schools as communities, poverty, levels of student populations, and students' attitudes, motives, and performance: A multilevel analysis. *American Educational Research Journal, 32*(3), 627-658.

Battistich, V., Solomon, D., Watson, M., & Schaps, E. (1997). Caring school communities. *Educational Psychologist, 32*(3), 137-151.

Baumeister, R., & Leary, M. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497-529.

Bertucci, A., Conte, S., Johnson, D., & Johnson, R. (2010). The impact of size of cooperative group on achievement, social support, and self-esteem. *The Journal of General Psychology, 137*(3), 256-272.

Bertucci, A., Johnson, D., Johnson, R., & Conte, S. (2011). The effects of task and resource interdependence on achievement and social support: An exploratory study of Italian children. *The Journal of Psychology, 145*(4), 343-360.

Bertucci, A., Johnson, D., Johnson, R., & Conte, S. (2012). Influence of group processing on acheivement and perception of social and academic support in elementary inexperienced cooperative learning groups. *The Journal of Educational Research, 105*(5), 329-335.

Bertucci, A., Johnson, D., Johnson, R., & Conte, S. (2012). Influence of group processing on achievement and perception of social and academic support in elementary inexperienced cooperative learning groups. *The Journal of Educational Research, 105*(5), 329-335.

Bierman, K. (2004). *Peer rejection: Developmental Processes and Intervention Strategies.* New York: Guilford Press.

Boland, A., Cherry, M., & Dickson, R. (2014). *Doing a Systematic Review: A Student's Guide.* London: SAGE Publcations.

Brace, N., Kemp, R., & Snelgar, R. (2012). *SPSS for Psychologists.* Basingstoke, Hampshire: Palgrave Macmillan.

British Psychological Society. (2009). *Code of Ethics and Conduct.* Leicester: British Psychological Society.

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

Brown, B., Mory, M., & Kinney, D. (1994). Casting adolescent crowds in a relational perspective: Caricature, channel and context. In R. Montemayor, G. Adams, & T. Gullotta, *Personal Relationships During Adolescence.* London: Sage.

Buchner, A., Erdfelder, E., Faul, F., & Lang, A. (2016, December 2nd). *G*Power version 3.1.9.2*. Retrieved from http://www.gpower.hhu.de/.

Buhs, E., & Ladd, G. (2001). Peer rejection as antecedent of young children's school adjustment: An examination of mediating processes. *Developmental Psychology, 37*, 550-560.

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

Bukowskil, W., & Hoza, B. (1989). Popularity and friendship issues in theory, measurement and outcome. In T. Berndt, & G. Ladd, *Peer Relations in Child Development* (pp. 15-45). New York: Wiley.

Butel, J., Braun, K., Novotny, R., Acosta, M., Castro, R., Fleming, T., . . . Nigg, C. (2015). Assessing intervention fidelity in a multi-level, multi-component, multi-site program: The children's healthy living (CHL) program. *Translational Behavioral Medicine, 5*(4), 460-469.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*(40). doi:10.1186/1748-5908-2-40

Cillessen, A., & Mayeux, L. (2004). Sociometric status and peer group behavior: Previous findings and current directions. In J. Kupersmidt, & K. Dodge, *Children's Peer Relations* (pp. 3-20). Washington DC: American Psychological Association.

Cohen, E. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research, 64*, 1-35.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education.* London: Taylor and Francis.

Coie, J., Dodge, K., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age persepective. *Developmental Psychology, 18*, 557-570.

Coie, J., Dodge, K., & Kupersmidt, J. (1990). Peer group behaviour and social status. In S. Asher, & J. Coie, *Peer Rejection in Childhood* (pp. 17-59). Cambridge: Cambridge University Press.

Coleman, J., & Hendry, L. (1999). *The Nature of Adolescence.* London: Routledge.

Collins. (2017, Feb 27th). Collins English Dictionary. Retrieved from https://www.collinsdictionary.com/dictionary/english

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

Craig, J. (2010). *Promoting peer acceptance in the classroom: An evaluation of a cooperative learning intervention in a mainstream primary school. Unpublished doctoral thesis.* Nottingham: The University of Nottingham.

Crick, N., & Dodge, K. (1994). A review and reformulation of social information procressing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74-101.

Crick, N., & Ladd, G. (1990). Children's perceptions of the outcomes of social strategies: Do the ends justify being mean? *Developmental Psychology, 4*, 612-620.

Dancey, C., & Reidy, J. (2014). *Statistics Without Maths for Psychology.* Chelmsford: Pearson Education Limited.

Dansereau, D., & Johnson, D. (1994). Cooperative Learning. In D. Druckman, & R. Bjork, *Learning, Remembering, Believing: Enhancing Human Performance* (pp. 83-111). Washington, DC: National Academy Press.

Demaray, M., & Elliott, S. (2001). Perceived social support by children with characteristics of attention-deficit/hyperactivity disorder. *School Psychology Quarterly, 16*, 68-90.

Demaray, M., & Malecki, C. (2002). The relationship between perceived social support and maladjustment for students at risk. *Psychology in the Schools, 39*(3), 305-316.

Demaray, M., Malecki, C., Davidson, L., Hodgson, K., & Rebus, J. (2005). The relationship between social support and student adjustment: A longitudinal analysis. *Psychology in the Schools, 42*(7), 691-706.

DeRosier, M., Kupersmidt, J., & Patterson, C. (1994). Children's academic and behavioral adjustment as a function of the chronicity and proximity of peer rejection. *Child Development, 65*(6), 1799-1813.

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

Dewey, J. (1966). *Democracy and Education.* New York: The Free Press.

Diehl, D., Lemerise, E., Caverly, S., Ramsay, S., & Roberts, J. (1998). Peer relations and school adjustment in ungraded primary children. *Journal of Educational Psychology, 90*(3), 506-515.

Dimitrov, D., & Rumrill, P. (2003). Pretest-posttest designs and measurement of change. *Work, 20*, 159-163.

Division of Educational and Child Psychology. (2002). *Professional Practice Guidelines.* Leicester: British Psychological Society.

Dodge, K., Lansford, J., Burks, V., Bates, J., Pettit, G., Fontaine, R., & Price, J. (2003). Peer rejection and social information processing factors in the development of aggressive behavior problems in children. *Child Development, 74*(2), 374-393.

Dodge, K., Pettit, G., McClaskey, C., Brown, M., & Gottman, J. (1986). Social competence in children. *Monographs of the Society for Research in Child Development, 51*(2), 1-85.

Downs, S., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non randomised studies of health care interventions. *Journal of Epidemiology Community Health, 52*, 377-384.

Doyle, L., Brady, A., & Byrne, G. (2009). An overview of mixed methods research. *Journal of Research in Nursing, 14*(2), 175-185.

East, P., Hess, L., & Lerner, R. (1987). Peer social support and adjustment of early adolescent peer groups. *Journal of Early Adolescence, 7*(2), 153-163.

Erikson, F., & Gutierrez, K. (2002). Culture, rigor, and science in educational research. *Educational Researcher, 31*(8), 21-24.

Erwin, P. (1998). *Friendship in Childhood and Adolescence.* London: Routledge.

Feilzer, Y. M. (2010). Doing mixed methods research pragmatically: Implications for the rediscovery of pragmatism as a research paradigm. *Journal of Mixed Methods Research, 4*(1), 6-16.

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics.* London: Sage Publications Ltd.

Frederickson, N. (1991). Children can be so cruel: Helping the rejected child. In G. Lindsay, & A. Miller, *Psychological Services for Primary Schools.* Harlow, Essex: Longman.

Frederickson, N. L., & Furnham, A. (1998b). Sociometric classification methods in school peer groups: A comparative investigation. *Journal of Clinical and Child Psychology, 6*, 921-934.

Frederickson, N., & Cline, T. (2009). *Special Educational Needs: Inclusion and Diversity.* Maidenhead : Open University Press.

Frederickson, N., & Furnham, A. (1998). *Social Inclusion Survey (SIS). Psychology in Education Portfolio: Social Skills and Emotional Intelligence.* NFER-Nelson.

Frederickson, N., Warren, L., & Turner, J. (2005). Circle of Friends - An exploration of impact over time. *Educational Psychology in Practice, 21*(3), 197-217.

Furrer, C., & Skinner, E. (2003). Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology, 95*, 148-162.

Gearing, R., El-Bassel, N., Ghesquiere, A., Baldwin, S., Gillies, J., & Ngeow, E. (2011). Major ingredients of fidelity: A review and scientific guide to

improving quality of intervention research implementation. *Clinical Psychology Review, 31*, 79-88.

Gersten, R., Fuchs, L., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Council for Exceptional Children, 71*(2), 149-164.

Gest, S., Graham-Bermann, S., & Hartup, W. (2001). Peer experience: Common and unique features of number of friendships, social network centrality, and sociometric status. *Social Development, 10*, 23-40.

Ghaith, G. (2003). The relationship between forms of instruction, achievement and perceptions of classroom climate. *Educational Research, 45*(1), 83-93.

Gifford-Smith, M., & Brownell, C. (2003). Childhood peer relationships: social acceptance, friendships, and peer networks. *Journal of School Psychology, 41*, 235-284.

Gillies, R. (2006). Teachers' and students' verbal behaviours during cooperative and small-group learning. *British Journal of Educational Psychology, 76*(2), 271-287.

Gillies, R. (2015). *Enhancing Classroom-Based Talk: Blending Practice, Research and Theory.* London: Routledge.

Gillies, R. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education, 41*(3), 39-54.

Gillies, R., & Ashman, A. (1995). The effects of gender and ability on students' behaviours and interactions in classroom based work groups. *British Journal of Educational Psychology, 65*, 211-225.

Gillies, R., & Ashman, A. (1997). The effects of training in cooperative learning on differential student behavior and achievement. *Journal of Classroom Interaction, 32*(1), 1-10.

Gillies, R., & Ashman, A. (2000). The effects of cooperative learning on students with learning difficulties in the lower elementary school. *Journal of Special Education, 34*, 19-27.

Gillies, R., & Ashman, A. (2003). An historical review of the use of groups to promote socialization and learning. In R. Gillies, & A. Ashman, *Co-Operative Learning: The Social and Intellectual Outcomes of Learning in Groups* (pp. 1-18). London: RoutledgeFalmer.

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

Gough, D., Oliver, S., & Thomas, J. (2013). *An Introduction to Systematic Reviews.* London: SAGE Publications.

Graves, N., & Graves, T. (1990). *A Part to Play: Tips, Techniques and Tools for Learning Cooperatively.* Victoria, Australia: Latitude Publications.

Gray, D. (2014). *Doing Research in the Real World.* London: Sage.

Gresham, F., & Stuart, D. (1992). Stability of sociometric assessment: Implications for uses as selection and outcome measures in social skills training. *The Journal of School Psychology, 30*, 223-231.

Griffiths, R., Stenner, R., & Hicks, U. (2014). Hearing the unheard: Children's constructions of their nurture group experiences. *Educational and Child Psychology, 31*(1), 124-136.

Grissom, R., & Kim, J. (2011). *Effect Sizes for Research: Univariate and Multivariate Applications.* New York: Routledge.

Gronlund, N. (1981). *Measurement and Evaluation in Teaching.* New York: Collier-Macmillan.

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

Harcum, E. (1990). Guidance from the literature for accepting a null hypothesis when its truth is expected. *Journal of General Psychology, 117*(3), 325-344.

Harden, A., & Gough, D. (2012). Quality and relevance appraisal. In D. O. Gough, & J. Thomas, *An Introduction to Systematic Reviews* (pp. 153-178). London: Sage Publications.

Harris, A., McGregor, J., Perencevich, E., Furuno, J., Zhu, J., Peterson, D., & Finkelstein, J. (2006). The use and interpretation of quasi experimental studies in medical informatics. *Journal of American Medical Informatics Association, 13*(1), 16-23.

Harrist, A., & Bradley, K. (2003). 'You can't say you can't play': Intervening in the process of social exclusion in the kindergarten classroom . *Early Childhood Research Quarterly*, 185-205.

Harter, S. (1996). Teacher and classmate influences on scholastic motivation, self esteem, and level of voice in adolescents. In J. Juvonen, & K. Wentzel, *Social Motivation: Understanding Children's School Adjustment* (pp. 11-42). Cambridge: Cambridge University Press.

Hartup, W. (1989). Social relationships and their developmental significance. *American Psychologist, 44*(2), 120-126.

Hattie, J. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement.* Routledge: Oxon.

Hay, D. (2005). Early peer relations and their impact on children's development. In R. Tremblay, R. Barr, & R. Peters, *Encyclopedia on Early Childhood Development* (pp. 1-6). Montreal, Quebec: Centre of Excellence for Early Childhood Development.

Health and Care Professions Council. (2012). *Standards of Conduct, Performance and Ethics.* London: Health and Care Professions Council.

Healy, M., & Perry, C. (2000). Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm. *Qualitative Market Research: An International Journal, 3*(3), 118-126.

Hendrickx, M., Mainhard, M., Boor-Klip, H., Cillessen, A., & Brekelmans, M. (2016). Social dynamics in the classroom: Teacher support and conflict and the peer ecology. *Teaching and Teacher Education, 53*, 30-40.

Herbert, M. (2002). *Working with Children, Adolescents and their Families.* Malden, MA: British Psychological Society.

Hertz-Lazarowitz, R., & Shachar, H. (1990). Teachers verbal behavior in cooperative and whole class instruction. In S. Sharan, *Cooperative Learning: Theory and Research* (pp. 77-94). New York: Praeger.

Higgins, S., Kokotsaki, D., & Coe, R. (2012). *The Sutton Trust - Education Endowment Foundation Teaching and Learning Toolkit: Technical Appendices.* London: Education Endowment Foundation.

Hinton, P. (2014). *Statistics Explained.* Hove, East Sussex: Routledge.

Hirsch, B., & DuBois, D. (1992). The relation of peer social support and psychological symptomatology during the transition to junior high school: A two year longitundinal analysis. *American Journal of Community Psychology, 20*, 333-347.

Hughes, J., Cavell, T., & Willson, V. (2001). Further support for the developmental significance of the quality of the teacher-student relationship. *Journal of School Psychology, 39*, 289-301.

Hymel, S., Wagner, E., & Butler, L. (1990). Reputational bias: View from the peer group. In S. Asher, & J. Coie, *Peer Rejection in Childhood* (pp. 156-186). Cambridge: Cambridge University Press.

Jacobs, D., Watson, T., Sutton, J., & Jones, B. (1996). Effects of a cooperative learning method on mathematics achievement and affective outcomes of students in a private elementary school. *Journal of Research and Development in Education, 29*(4), 195-202.

Jacques, N., Wilton, K., & Townsend, M. (1998). Cooperative learning and social acceptance of children with mild intellectual disability. *Journal of Intellectual Disability Research, 42*(1), 29-36.

Johnson, D., & Johnson, F. (2000). *Joining Together: Group Theory and Group Skills.* Boston: Allyn and Bacon.

Johnson, D., & Johnson, R. (1999). *Learning Together and Alone: Cooperative, Competitive and Individualistic Learning.* Boston: Allyn and Bacon.

Johnson, D., & Johnson, R. (2002). Learning together and alone: Overview and meta analysis. *Asia Pacfic Journal of Education, 22*, 95-105.

Johnson, D., & Johnson, R. (2005). New developments in social interdependence theory. *Psychology Monographs, 131*, 285-358.

Johnson, D., & Johnson, R. (2008). Social interdependence theory and cooperative learning: The teacher's role. In R. Gillies, A. Ashman, & J. Terwel, *The Teacher's Role in Implementing Cooperative Learning in the Classroom* (pp. 9-37). New York: Springer.

Johnson, D., & Johnson, R. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*, 365-379.

Johnson, D., & Johnson, R. (2013). Cooperative, competitive, and individualistic learning environments. In J. Hattie, & E. Anderman, *International Handbook of Student Achievement* (pp. 372-374). New York: Routledge.

Johnson, D., Johnson, F., & Stanne, M. (2000). Cooperative Learning Methods: A Meta Analysis. Retrieved March 1st, 2017, from https://pdfs.semanticscholar.org/93e9/97fd0e883cf7cceb3b1b612096c27aa40f90.pdf

Johnson, D., Johnson, R., & Anderson, D. (1983). Social interdependence and classroom climate. *The Journal of Psychology, 114*, 135-142.

Johnson, D., Johnson, R., & Holubec, R. (1994). *The New Circles of Learning: Cooperation in the Classroom and School.* ASCD: Alexandria, Virginia.

Johnson, D., Johnson, R., & Taylor, B. (1993). Impact of cooperative and individualistic learning on high ability students' achievement, self-esteem, and social acceptance. *The Journal of Social Psychology, 133*(6), 839-844.

Johnson, D., Johnson, R., Buckman, L., & Richards, P. (1985). The effect of prolonged implementation of cooperative learning on social support within the classroom. *The Journal of Psychology, 119*(5), 405-411.

Johnson, D., Johnson, R., Stanne, M., & Garibaldi, A. (1990). Impact of group processing on achievement in cooperative groups. *Journal of Social Psychology, 130*, 507-516.

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

Jolliffe, W. (2007). *Cooperative Learning in the Classroom: Putting it into Practice.* London: Paul Chapman Publishing.

Jordan, D., & Le Metais, J. (1997). Social skilling through cooperative learning. *Educational Research, 39*(1), 3-21.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development.* Alexandria, VA: Association for Supervision and Curriculum Development.

Ju, E. (2009). Is Random Assignment Good Enough? *Journal of Policy Analysis and Management, 28*(1), 176-178.

Jules, V. (1991). Interaction dynamics of cooperative learning groups in Trinidad's secondary schools. *Adolescence, 26*(104), 931-931.

Juvonen, J. (2007). Reforming middle schools: Focus on continuity, social connectedness, and engagement. *Educational Psychologist, 42*(4), 197-208.

Kagan, S. (1990). The structural approach to cooperative learning. *Educational Leadership, 47*(4), 12-15.

Kagan, S., & Kagan, M. (1994). The structural approach: Six keys to cooperative learning. In S. Sharan, *Handbook of Cooperative Learning Methods* (pp. 115-133). Westport, CT: Greenwood.

Kagan, S., & Kagan, M. (2009). *Kagan Cooperative Learning.* San Clemente, CA: Kagan Publishing.

Kirk, R. (1999). *Statistics: An Introduction.* London: Harcourt Brace.

Kirk, T. (1997). *The effectiveness of cooperative learning: With particular reference to academic achievement, self esteem, academic self image, social interaction and student attitudes in primary mathematics and english spelling classes in Ireland. .* Unpublished doctoral thesis. Dublin: University of Dublin.

Knapp, T., & Schafer, W. (2009). From gain score t to ANCOVA F (and vice versa). *Practical Assessment, Research and Evaluation, 14*(6), 1-7.

Kupersmidt, J., & Coie, J. (1990). Preadolescent peer status, aggression, and school adjustment as predictors of externalising problems in adolescence. *Child Development, 61*, 1350-1362.

Ladd, G. (1983). Social networks of popular, average and rejected children in social settings. *Merrill Palmer Quarterly, 29*, 282-307.

Ladd, G. (2006). Peer rejection, aggressive or withdrawn behavior, and psychological maladjustment from ages 5 to 12: An examination of four predictive models. *Child Development, 77*(4), 822-846.

Ladd, G., & Asher, S. (1985). Social skill training and children's peer relations: Current issues in research and practice. In L. L'Abate, & M. Milan, *Handbook of Social Skill Training* (pp. 219-244). New York: Wiley.

Ladd, G., & Burgess, K. (2001). Do relational risks and protective factors moderate the linkages between childhood aggression and early psychological and school adjustment? *Child Development, 72*, 1579-1601.

Ladd, G., & Troop-Gordon, W. (2003). The role of chronic peer difficulties in the development of children's psychological adjustment problems. *Child Development, 74*(5), 1344-1367.

LaFont, L., Proeres, M., & Vallet, C. (2007). Cooperative group learning in a team game: Role of verbal exchanges among peers. *Social Psychology of Education, 10*, 93-113.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t tests and ANOVAs. *Frontiers in Psychology, 4*, 1-12.

Lakey, B., & Cohen, S. (2000). Social support theory and measurement. In S. Cohen, L. Underwood, & B. Gottlieb, *Measuring and Intervening in Social Support* (pp. 29-52). New York: Oxford University Press.

Lanier, S. (2007). *Multiple Aspects of Children's Perceptions of Classroom Peer Support and Adjustment.* University of Maryland: ProQuest Dissertations.

Lazarowitz, R., Hertz-Lazarowtiz, R., & Hugh, B. (1994). Learning science in a cooperative setting: Academic achievement and affective outcomes. *Journal of Research in Science Teaching, 31*(10), 1121-1131.

Levett-Jones, T., Lathlean, J., Maguire, J., & McMillan, M. (2007). Belongingness: A critique of the concept and implications for nursing education. *Nurse Education Today, 27*(3), 210-218.

Lewin, K. (1935). *A Dynamic Theory of Personality.* New York: McGraw-Hill.

Lou, Y., Abrami, P., Spence, J., Poulsen, C., Chambers, B., & d'Apollonia, S. (1996). Within-class grouping: A meta analysis. *Review of Educational Research, 66*, 423-458.

MacMillan, D., & Morrison, G. (1984). Sociometric research in special education. In R. Jones, *Attitudes and Attitude Change in Special Education: Theory and Practice* (pp. 93-117). Reston, VA: Council for Exceptional Children.

Malecki, C., & Demaray, M. (2003). Carrying a weapon to school and perceptions of social support in an urban middle school. *Journal of Emotional and Behavioral Disorders, 11*, 169-178.

Malecki, C., & Elliott, S. (1999). Adolescents' ratings of perceived social support and its importance: validation of the student support scale. *Psychology in the Schools, 36*(6), 473-482.

Marr, M. (1997). Cooperative learning: A brief review. *Reading and Writing Quarterly, 13*(1), 7-20.

Maslow, A. (1970). *Motivation and Personality.* New York: Harper and Row.

McKown, C., Gumbiner, L., & Johnson, J. (2011). Diagnostic efficiency of several methods of identifying socially rejected children and effect of participation rate on classfication accuracy. *Journal of School Psychology, 49*, 573-595.

Mercer, N., & Littleton, K. (2007). *Dialogue and the Development of Children's Thinking: A Sociocultural Approach.* New York: Routledge.

Mertens, D. (2015). *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods.* London: SAGE Publishing.

Mikami, A., Lerner, M., & Lun, J. (2010). Social context influences on children's rejection by their peers. *Child Development Perspectives, 4*(2), 123-130.

Miller, G., & Chapman, J. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*(10), 40-48.

Morgan, D. (2007). Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research, 1*, 48-76.

Muijs, D. (2010). *Doing Quantitative Research in Education with SPSS.* London: SAGE Publications.

Murphy, E., Grey, I., & Honan, R. (2005). Cooperative learning for students with difficulties in learning: A description of models and guidelines for implementation. *British Journal of Special Education, 32*(3), 157-164.

Murphy, S., & Gutman, S. (2012). Intervention fidelity: A necessary aspect of intervention effectiveness studies. *American Journal of Occupational Therapy, 66*, 387-388.

Nansel, T., Overpeck, M., Pilla, R., Ruan, W., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors amoung US Youth: Prevalence and association with psychosocial adjustment. *Journal of the American Medical Association, 285*(16), 2094-2100.

Nelson, L., Rubin, K., & Fox, N. (2005). Social and non social behaviors and peer acceptance: A longitudinal model of the development of self-perceptions in children ages 4 to 7 years. *Early Education and Development, 20*, 185-200.

Newcomb, A., Bukowski, W., & Pattee, L. (1993). Children's peer relations: a meta-analytic review of popular, rejected, neglected, controversial and average sociometric status. *Psychological Bulletin, 113*(1), 99-128.

Nock, M., Michel, B., & Photos, V. (2007). Single-Case Research Designs. In D. Mckay, *Handbook of Research Methods in Abnormal and Clinical Psychology* (pp. 337-350). Thousand Oaks, CA: Sage Publications.

Noll, R., Zeller, M., Vannatta, K., Bukowski, W., & Davies, W. (1997). Potential bias in classroom research: Comparison of children with permission and those who do not receive permission to participate. *Journal of Clinical Child Psychology, 26*, 36-42.

O'Connor, R., & Jenkins, J. (2013). Cooperative learning for students with learning disabilities: Advice and caution derived from the evidence. In H. Swanson, K. Harris, & S. Graham, *Handbook of Learning Disabilities* (pp. 507-525). New York: The Guilford Press.

Oortwijn, M., Boekaerts, M., Vedder, P., & Fortuin, J. (2008). The impact of a cooperative learning experience on pupils' popularity, non-cooperativeness, and interethnic bias in multiethnic elementary schools. *Educational Psychology, 28*(2), 211-221.

Ottenbacher, K. (1998). Quantitative evaluation of multiplicity in epidemiology and public health research. *American Journal of Epidemiology, 147*(7), 615-619.

Parker, J., & Asher, S. (1987). Peer relations and later personal adjustment: Are low-accepted chidlren at risk? *Psychological Bulletin, 102*(3), 357-389.

Peterson, R., Johnson, D., & Johnson, R. (1991). Effects of cooperative learning on perceived status of male and female pupils. *The Journal of Social Psychology, 131*(5), 717-735.

Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide.* Oxford: Blackwell.

Piaget, J. (1959). *The Language and Thought of the Child.* London: Routledge.

Piercy, M., Wilton, K., & Townsend, M. (2002). Promoting the social acceptance of young children with moderate-severe intellectual disabilities using cooperative-learning techniques. *American Journal on Mental Retardation, 107*(5), 352-360.

Pisula, E., & Lukowska, E. (2011). Perception of social relationships with classmates and social support in adolescents with Asperger syndrome attending mainstream schools in Poland. *School Psychology International, 33*(2), 185-206.

Price, J., & Dodge, K. (1989). Peers contributions to children's social maladjustment. In T. Berndt, & G. Ladd, *Peer Relationships in Child Development* (pp. 341-370). Chichester: John Wiley & Sons.

Putnam, J., Markovchick, K., Johnson, D., & Johnson, R. (1996). Cooperative learning and peer acceptance of students with learning disabilities. *The Journal of Social Psychology*, 741-752.

Reilly, R., & Mitchell, S. (2010). The clash of the paradigms: Tracking, cooperative learning, and the demolition of the zone of proximal development. *The Alberta Journal of Educational Research, 56*(4), 419-435.

Renk, K., & Phares, V. (2004). Cross-informant ratings of social competence in children and adolescents. *Clinical Psychology Review, 24*(2), 239-254.

Roberts, W., Horn, A., & Battistich, V. (1995). Assessing students' and teachers' sense of the school as a caring community. *Annual Meeting of the American Educational Research Association.* San Francisco, California: www.devstu.org.

Robson. (2011). *Real World Research.* Chichester: Wiley-Blackwell.

Roeser, R., Eccles, J., & Sameroff, A. (1998). Academic and emotional functioning in early adolescence: Longitundinal relations, patterns, and prediction by experience in middle school. *Development and Psychopathology, 7*, 295-312.

Rosenbaum, P. (1987). The role of a second control group in an observational study. *Statistical Science, 2*, 292-316.

Roseth, C., Johnson, D., & Johnson, R. (2008). Promoting early adolescents' achievement and peer relationships: The effects of cooperative, competitive, and indivdualistic goal structures. *Psychological Bulletin, 134*, 223-246.

Rothman, K. (1990). No adjustments are needed for multiple comparisons. *Epidemiology, 1*(1), 43-46.

Rubin, K., Begle, A., & McDonald, K. (2012). Peer relations and social competence in childhood. In V. Anderson, & M. Beauchamp, *Developmental Social Neuroscience and Childhood Brain Insult: Implications for Theory and Practice* (p. Chapter 2). New York: Guilford Press. Retrieved from https://www.researchgate.net/publication/234026355_Peer_Relations_and_Social_Competence_in_Childhood

Rzoska, K., & Ward, C. (1991). The effects of cooperative and competitive learning methods on the mathematics achievement, attitudes toward school, self-concepts and friendship choices of Maori, Pakeha and Samoan children. *New Zealand Journal of Psychology, 20*, 17-24.

Saylor, C., & Leach, J. (2009). Perceived bullying and social support in students accessing special inclusion programming. *Journal of Developmental and Physical Disabilities, 21*, 69-80.

Schagen, I., & Hodgen, E. (2009). *How Much Difference Does it Make? Notes on Understanding, Using, and Calculating Effect Sizes for Schools.* Wellington, New Zealand: Research Division, Ministry of Education.

Schmidt, F., & Hunter, J. (2002). Are there benefits from NHST? *American Psychologist, 57*(1), 65-66.

Scotland, J. (2012). Exploring the philospohical underpinnings of research: relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching, 5*(9), 9-16.

Sen, A. (2000). *Social Exclusion: Concept, Application, and Scrutiny.* Manila, Philippines: Asian Development Bank.

Senn, S. (2006). Change from baseline and analysis of covariance revisited. *Statistics in Medicine, 25*, 4334-4344.

Shachar, H. (2003). Who gains what from co-operative learning: An overview of eight studies. In R. Gillies, & A. Ashman, *Co-Operative Learning: The*

Social and Intellectual Outcomes of Learning in Groups (pp. 103-935). London: RoutledgeFalmer.

Sheldon, S. (2004). Social relationships and health. *American Psychologist*, 676-684.

Silver, H. (2007). *Social Exclusion: Comparative Analysis of Europe and Middle East Youth.* Middle East Youth Initiative Working Paper.

Skidmore, D., & Gallagher, D. (2005). A Dialogical Pedagogy for Inclusive Education. *Inclusive and Supportive Education Congress.* Glasgow, Scotland.

Skinner, B. (1953). *Science and Human Behaviour.* New York: Macmillan.

Slater, A., & Bremner, G. (2011). *An Introduction to Developmental Psychology.* Oxford: Wiley, John and Sons.

Slavin, R. (1980). Cooperative learning. *Review of Educational Research, 50*(2), 315-342.

Slavin, R. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin, 94*, 429-445.

Slavin, R. (1989). Cooperative learning and student achievement. In R. Slavin, *School and Classroom Organization* (pp. 129-156). New Jersey: Lawrence Erlbaum.

Slavin, R. (1995). *Cooperative Learning.* Boston, USA: Allyn and Bacon.

Slavin, R. (2004). When and why does cooperative learning increase achievement? . In H. Daniels, & A. Edwards, *The RoutledgeFalmer Reader in Psychology of Education* (pp. 271-293). London: RoutledgeFalmer.

Slavin, R. (2013). Effective programmes in reading and mathematics: lessons from the best evidence encyclopaedia. *School Effectiveness and School Improvement, 24*(4), 383-391.

Slavin, R. (2014). Cooperative learning and academic achievement: Why does groupwork work? *Anales de Psicologia, 30*, 784-791.

Smith, P., Cowie, H., & Blades, M. (2011). *Understanding Children's Development.* Chichester: John Wiley & Sons Ltd.

Streiner, D., & Norman, G. (2011). Correction for multiple testing: Is there a resolution? *Chest, 140*, 16-18.

Symes, W., & Humphrey, N. (2010). Peer-group indicators of social inclusion among pupils with autistic spectrum disorders (ASD) in mainstream secondary schools: A comparative study. *School Psychology International, 31*(5), 478-494.

Terry, R., & Coie, J. (1991). A comparison of methods for defining sociometric status among children. *Developmental Psychology, 27*(5), 867-880.

The University of Nottingham. (2015). *Code of Research Conduct and Research Ethics .* Nottingham: The University of Nottingham.

Thomson, C., & Brown, D. (2000). Co-operative learning: What it has to offer New Zealand teachers. *Learning in Partnership, 3*, 38-41.

Thurston, A., Topping, K., Tolmie, A., Christie, D., Karagiannidou, E., & Murray, P. (2010). Cooperative learning in science: Follow up from primary to high school. *International Journal of Science Education, 32*(4), 501-522.

Toner, M., & Munro, D. (1996). Peer-social attributions and self-efficacy of peer-rejected preadolescents. *Merrill-Palmer Quarterly, 42*(3), 339-357.

Topping, K., Thurston, A., Tolmie, A., Christie, D., Murray, P., & Karagiannidou, E. (2011). Cooperative learning in science: Intervention in the secondary school. *Research in Science and Technological Education, 29*(1), 91-106.

Trochim, W., Donnelly, J., & Arora, K. (2016). *Research Methods: The Essential Knowledge Base.* Boston: Cengage.

Vasileiadou, M. (2009). Cooperative learning and its effects on pre-primary, marginalized children. *Emotional and Behavioural Difficulties, 14*(4), 337-347.

Vygotsky, L. (1978). *Mind in Society: The Development of Higher Psychological Processes.* Cambridge, MA: Harvard University Press.

Waas, G., & Honer, S. (1990). Situational attributions and dispositional influences: the development of peer reputation. *Merrill Palmer Quarterly, 2*, 239-260.

Webb, N. (1985). Student interaction and learning in small groups: A research summary. In R. Slavin, S. Sharon, S. Kagan, R. Hertz-Larowitz, C. Webb, & R. Schmuck, *Learning to Cooperate, Cooperating to Learn* (pp. 5-15). New York: Plenum.

Webb, N., & Cullican, L. (1983). Group interactions and achievement in small groups: Stability over time. *American Educational Research Journal, 20*, 411-423.

Wentzel, K. (2003). Sociometric status and adjustment in middle school: A longitudinal study. *Journal of Early Adolescence, 23*, 5-28.

Wentzel, K., & Caldwell, K. (1997). Friendships, peer acceptance, and group membership: Relations to academic achievement in middle school. *Child Development, 68*(6), 1198-1209.

Wenz-Gross, M., & Siperstein, G. (1998). Students with learning problems at risk in middle school: Stress, social support, and adjustment. *Exceptional Children, 65*, 91-100.

Wenz-Gross, M., Siperstein, G., & Untch, A. (1997). Stress, social support, and adjustment of adolescents in middle schools. *Journal of Early Adolescence, 17*, 129-151.

Wethington, E., & Kessler, R. (1986). Perceieved support, received support, and adjustment to stressful life events. *Journal of Health and Social Behavior, 27*, 78-89.

Wildt, A., & Ahtola, O. (1978). *Analysis of Covariance.* Beverly Hills, CA: Sage.

Williams, P., Barclay, L., & Schmied, V. (2004). Defining social support in context: A necessary step in improving research, intervention, and practice. *Qualitative Health Research, 14*(7), 942-960.

Wittrock, M. (1989). Generative processes of comprehension. *Educational Psychologist, 24*(4), 345-376.

Woodward, L., & Fergusson, D. (2000). Childhood peer relationship problems and later risks of educational under-acheivement and unemployment. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 2*, 191-201.

Wright, D. (2006). The art of statistics: A survey of modern statistics. In P. Alexander, & P. Winne, *Handbook of Educational Psychology* (pp. 879-901). Mahwah, New Jersey: Erlbaum.

Zakriski, A., & Coie, J. (1996). A comparison of aggressive-rejected and nonaggressive-rejected children's interpretations of self-directed and other-directed rejection. *Child Development, 67*, 1048-1070.

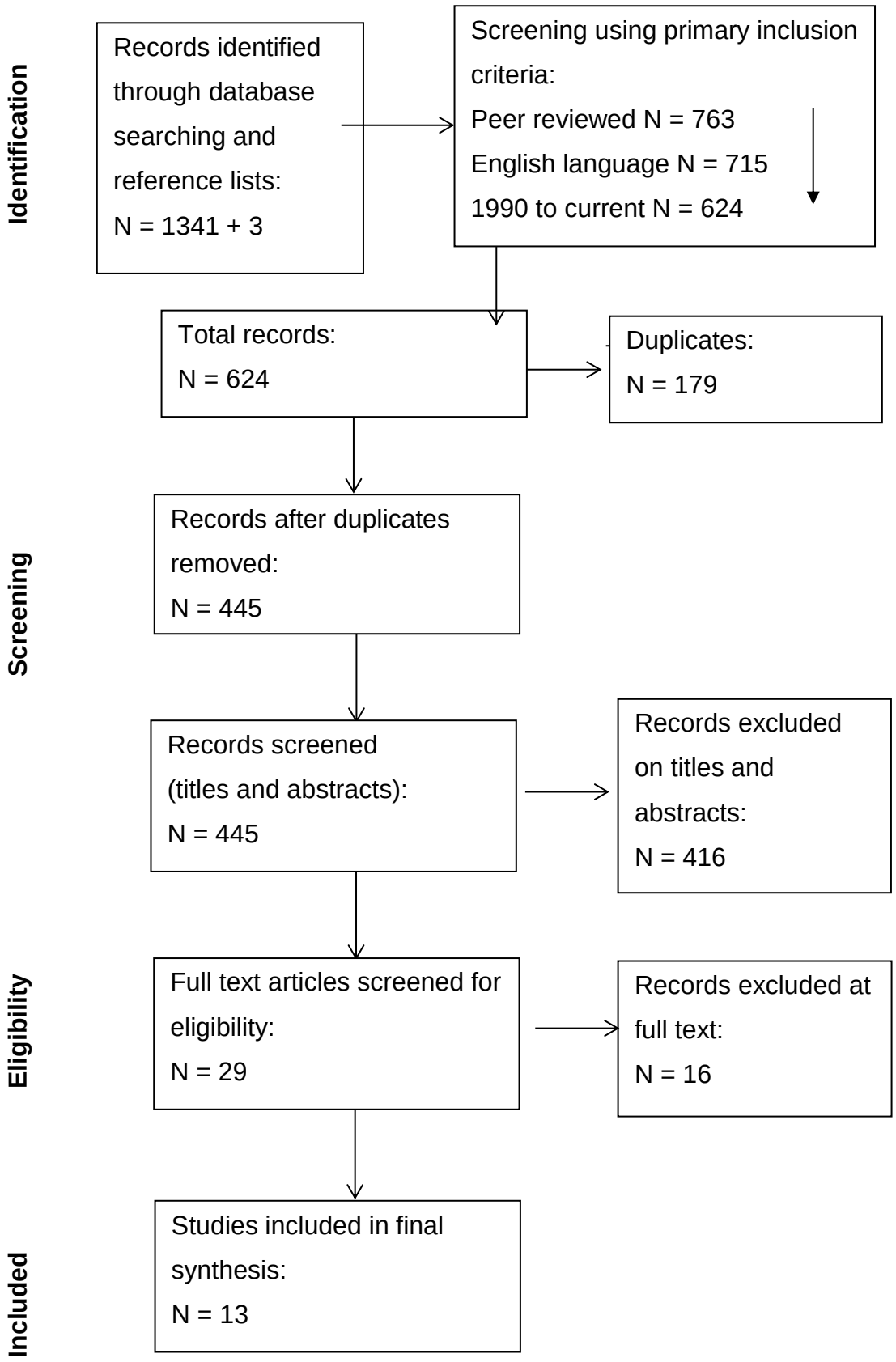
Zettergren, P. (2005). Childhood peer status as predictor of midadolescene peer situation and social adjustment. *Psychology in the Schools, 42*(7), 745-757.

Secondary References

Garel, J. (2001). L'intégration des élèves de SEGPA en Éducation physique et sportive: entre espoirs et illusions. *La Nouvelle Revue de l'AIS, 14,* 61-71.

6. Appendices

Appendix A A flow diagram of the study selection process



Inclusion criteria were developed with regard to the review question, as follows:

Primary inclusion criteria:

1. The study was published in a peer reviewed paper
2. The full paper was available in English
3. The study was published between 1990 and January 29th 2017

Secondary inclusion criteria:

4. The study evaluated the impact of a CL intervention with regards to review relevant outcomes
5. The study employed an experimental design, with repeated measures
6. The intervention took place in a school setting with a school based sample, and with participants aged between four and sixteen years old

Inclusion criteria were developed in order to increase the likelihood that the studies would be of relatively high quality, and allow an evaluation of the impact of CL according to the review question, in a population of primary interest and relevance to an EP.

After primary inclusion criteria were applied, there were 624 records remaining. Of these, 179 were duplicates. The secondary inclusion criteria were applied to the abstracts of 445 papers.

At the abstract level, papers were most frequently excluded due to the intervention not being CL, outcomes being unrelated to the focus of the review question (e.g. academic or teacher focused outcomes), participants being aged over 16, or the paper being a descriptive review of CL, or of a project which involved CL.

The full text of 29 papers which met the primary and secondary inclusion criteria were obtained. Of these, 16 papers were excluded for reasons such as only using post intervention measures, for outcomes which were not related to the review question, using participants over the age of 16, or not evaluating a CL

intervention (e.g. a descriptive case study, and/or no active CL intervention present).

Thirteen studies were suitable for inclusion.

Appendix B Full text studies excluded from synthesis

Reference	Reason for Exclusion
Battistich, V., Solomon, D. and Delucchi, K. (1993) Interaction processes and student outcomes in cooperative learning groups. *The Elementary School Journal,* 94, 1, 19-32.	Post intervention measures only. Did not answer the review question.
Gillies, R. (2004) The effects of cooperative learning on junior high students during small group learning. *Learning and Instruction,* 14, 197-213.	Did not answer the review question.
Gillies, R. and Ashman, A. (1997) The effects of training in cooperative learning on differential student behavior and achievement. *Journal of Classroom Interaction,* 32, 1, 1-10.	Did not answer the review question.
Gillies, R. and Ashman, A. (2000) The effects of cooperative learning on students with learning difficulties in the lower elementary school. *The Journal of Special Education*, 34, 1, 19-27.	Did not answer the review question.

Johnson, D., Johnson, R. and Taylor, B. (1993) Impact of cooperative and individualistic learning on high ability students' achievement, self-esteem, and social acceptance. *The Journal of Social Psychology,* 133, 6, 839-844.	Post intervention measures only.
Jordan, D. and Le Metais, J. (1997) Social skilling through cooperative learning. *Educational Research,* 39, 1, 3-21.	No pre and post measures. Descriptive case study,
Jules, V. (1991) Interaction dynamics of cooperative learning groups in Trinidad's secondary schools. *Adolescence,* 91, 26, 931-931.	Participants were largely over 16 years old.
Kadir, S., Luan, W., Pihie, Z., Yaacob, N., Tarmizi, R. and Elias, H. (2005) The effects of cooperative learning strategy on peer attachment. *Pakistan Journal of Psychological Research,* 20, 3-4, 121-131.	Participants were 16 – 20 years old. Post intervention measures only.
Klavina, A., Jerlinder, K., Kristen, L., Hammar, L. and Soulie, T. (2014) Cooperative orientated learning in inclusive physical education. *European Journal of Special Needs Education,* 29, 2, 119-134.	Did not answer the review question.

Lazarowitz, R., Hertz-Lazarowitz, R. and Hugh, B. (1994). Learning science in a cooperative setting: Academic achievement and affective outcomes. *Journal of Reserach in Science Teaching,* 31, 10, 1121-1131.	Participants were aged 16 – 18 years old.
Peterson, R., Johnson, D. and Johnson, R. (1991) Effects of cooperative learning on perceived status of male and female pupils. *The Journal of Social Psychology,* 131, 5, 717-735.	Did not answer the review question.
Reilly, R. and Mitchell, S. (2010) The clash of the paradigms: Tracking, cooperative learning, and the demolition of the zone of proximal development. *The Alberta Journal of Educational Research,* 56, 4, 419-435.	Did not answer the review question. Participants were aged 14 – 18 years old.
Rzoska, K. and Ward, C. (1991) The effects of cooperative and competitive learning methods on the mathematics achievement, attitudes towards school, self-concepts and friendship choices of Maori, Pakeha and Samoan children. *New Zealand Journal of Psychology,* 20, 17-24.	Post intervention measures only.
Shepperson, G. (1991) Cooperative learning in a silent classroom. *Reading Psychology,* 12, 2, 133-145.	Did not answer the review question. No pre and post measures. No clear Cooperative Learning intervention.

Thurston, A., Topping, K., Tolmie, A., Christie, D., Karagiannidou, E. and Murray, P. (2010) Cooperative learning in science: Follow up from primary to high school. *International Journal of Science Education,* 32, 4, 501-522.	No active Cooperative Learning intervention.
Thurston, A., Topping, K., Tolmie, A., Christie, D., Karagiannidou, E. and Murray, P. (2010) Enhancing outcomes in school science for pupils during transition from elementary school using cooperative learning. *Middle Grades Research Journal,* 5, 1, 19-32.	Duplicate of previous article (above). No active Cooperative Learning intervention.

Appendix C Key facts table providing details of each study included in the systematic review: Sample, design, and intervention

Study and Location	Sample	Design*	Intervention**
Andre, Deneuve and Louvet (2011) France	**Middle school** **6th grade** **Age: 11-12 years old** 217 total, including 32 children with academic learning difficulties (SEGPA). 105 intervention, 112 control. 122m, 95f 8 intact classes with pupils added from 4 SEGPA classes. 3m, 1f SEGPA added to each class 4 teachers, 4 schools	Random assignment to condition Cooperative Learning (experimental) Individualistic (control)	Johnson and Johnson CL 5 weeks (6 x 2 hour lessons) Approximate exposure** 12 hours PE (gymnastics) 2 days staff training Unknown social skills training

Andre, Louvet and Deneuve (2013) France	**Middle school** **Age: 11 – 12 years old** 168 pupils, including 36 focus pupils with academic learning difficulties (SEGPA) 95m, 73f 3 classes low risk, 3 classes high risk in intervention. 6 intact classes with pupils added from 3 SEGPA classes UK equivalent Year 7	Quasi-experimental 2 experimental groups only (CL low risk, CL high risk taking)	Adapted from Johnson and Johnson 7 x 2 hour lessons over 6 weeks Approximate exposure** 14 hours PE (gymnastics) Unknown staff training Unknown social skills training
Bertucci, Conte, Johnson and	**Italian 7th grade** 62 pupils (20 CL pairs, 24 CL fours, 18 individualistic)	Stratified random assignment	6 x 90 minute lessons over 6 weeks (weekly session) on alcohol, tobacco, and drug abuse

Johnson (2010) Italy	31m, 31f WISC-RL (Wechsler Intelligence Scale for Children, Revised) used to verify that no pupils had 'cognitive deficits'. 4 teachers rotated across conditions	2 intervention groups (CL pairs, CL fours) Individualistic (control)	Goal and task interdependence in CL conditions
Bertucci, Johnson, Johnson and Conte (2011) Italy	**Italian 7th grade** 66 pupils (24 task and goal CL, 24 resource and goal CL, 18 individualistic learning) 32m, 34f WISC-RL (Wechsler Intelligence Scale for Children, Revised) used to verify that no pupils had 'cognitive deficits'.	Stratified random assignment 2 intervention groups (CL positive task and goal interdependence; CL positive resource and	6 x 90 minute lessons on alcohol, tobacco, and drug abuse (45 minute lessons? inconsistent in text)

	3 classes	goal interdependence) Individualistic (control)	
Bertucci, Johnson, Johnson and Conte (2012) Italy	**Italian elementary school 3rd – 5th grade** 61 pupils (30 group processing, 31 no group processing) 29m, 32f	Random assignment to condition and work group 2 intervention groups only (CL with group processing, CL without group processing)	Goal interdependence 5 sessions x 90 minutes over a 3 week period Academic lessons (e.g. science, history, Italian) Researcher led

Jacobs et al. (1996) USA	**Private Christian elementary school 3rd to 5th grade** 266 pupils including 135 intervention (70m, 65f) and 131 control (69m, 62f) 6 classes, 1 school (3 classes intervention, 3 control)	Random assignment to condition Cooperative Learning (experimental) 'traditional instruction' (control)	Slavin STAD method 9 weeks (4 x 20 minutes weekly) Approximate exposure* 12 hours Maths 3 hour training and weekly ongoing support Unknown social skills training
Jacques, Wilton and Townsend (1998) New Zealand	**Age: 9 – 11 years old** 24 pupils with 'mild intellectual disability' from 21 schools (focus pupils). 12 usually attended mainstream class, 12 usually attended special class 21m, 3f	Random assignment to condition Cooperative Learning (experimental)	Based on Slavin (1983) jigsaw activities, and Aronson et al (1978). 6 weeks (4 x 30 minutes weekly) Approximate exposure** 12 hours Social studies / Geography

	Identified as WISC full scale IQ range 50-70.	'usual classroom' (control) 5 week follow up	No staff training – researcher led Unknown social skills training
LaFont, Proeres and Vallet (2007) France	**Primary school** **3rd and 4th grade** **Age: mean 9.4 years old** 30 pupils (15 intervention, 15 control) 17m, 13f. 3 teams intervention, 3 teams control	Random assignment to condition Cooperative Learning (experimental) Control (as usual)	Adapted from Johnson and Johnson 10 weeks (1 x 45 minutes each week) Approximate exposure** 7.5 hours PE (basketball) Unknown staff training
Oortwijn et al. (2008) Netherlands	**Elementary school** **5th grade** **Age: 10-12 years old**	Quasi-experimental	11 lessons x 1 hour (including 2 x 1hour social skills) Approximate exposure** 9 hours

	94 pupils (51m, 43f). 26 intervention teams	5 intervention classes	Maths
	5 classes. 5 elementary schools		2 hours staff training
		No control or comparison	Social skills training
Piercy, Wilton and Townsend (2002)	**2nd grade** **Age: 5-7 years old**	Random assignment to condition	Johnson and Johnson's (1999) learning together and alone
	51 pupils, including 6 male focus pupils with 'moderate to severe intellectual disability'	Cooperative	10 weeks (2 x 40 minutes weekly) Approximate exposure** 13 hours 20 minutes
New Zealand	(WISC 20-50 full scale IQ) who attended special classes	Learning (experimental)	Specially devised activities e.g. cutting and pasting, drawing and painting
	29m, 22f 1 school	Social contact (comparison)	Unknown staff training
	17 pupils in intervention, 17 social contact, 17 control.	Usual (control)	No social skills training

Putnam et al. (1996)	**5th – 8th grade** **Age: 10 – 15 years old**	Quasi-experimental	Johnson and Johnson 5 principles 8 months (2 x 45 minute lessons weekly)
USA	417 pupils (197m, 220f) + 41 pupils (32m, 9f) classified as 'disabled' by state guidelines e.g. learning disability, behavioural disorder. 2 schools, 21 classes, 17 teachers	Cooperative Learning (experimental) Competitive, individual (comparison)	Approximate exposure** 42-48 hours Academic subjects including maths, science, social studies and English 34 hours teacher training over 6 months (8 hours initial training, 5 monthly seminars, 2 day training). No social skills training
Topping et al. (2011)	**Secondary school** **Age: 12 – 14 years old**	Quasi-experimental	Frequency and lesson duration not specified. 5-6 months (including 1-2 months social skills training)
Scotland	259 intervention (127m, 132f), 385 comparison (181m, 203f) = 644 total	Cooperative Learning (experimental)	Science lessons

	9 secondary schools, 12 class teachers	Control (teaching as usual)	3 days staff training
Vasileiadou (2009)	**Pre-primary / Kindergarten Mean age: 5.8 years**	Intervention only	Johnson and Johnson 'learning together' method
Cyprus	13 children with consent (6m, 11f) in kindergarten (17 in the class)		Specially devised activities which required cooperation to accomplish successfully
	2 classes, 1 school. Mainly Cypriot children.		6 weeks with 3 intervention weeks (12 days) + 1 week social skills using cooperative games proposed by Hill (1992). 30 minute session each active day.
			No staff training - researcher delivered.

*** All studies involved pre-test, post-test procedures.**

** Approximate exposure was calculated by the researcher of the current paper and not taken from individual research papers. For Putnam et al. (1996) this calculation is particularly approximate.

Appendix D Key facts table providing details of each study included in the systematic review: Measures, dependent variables, and relevant findings

Study	Measure(s) used	Outcome(s) of interest (DV)	Relevant findings
Andre, Deneuve and Louvet (2011)	Frederickson and Furnham (1998) Social Inclusion Survey How much would you like to play sport with…?	Social preference and social impact	Significant increase pre to post-test for social preference of SEGPA pupils in CL condition, but no significant change in control condition. No significant change pre to post-test for social impact of SEGPA or mainstream pupils No significant change pre to post-test for social preference or social impact of mainstream pupils in either the intervention or control conditions.
Andre, Louvet and Deneuve (2013)	Frederickson and Furnham (1998) Social Inclusion Survey	Social preference and social impact	Significant increase pre to post-test for social preference of SEGPA pupils in high risk CL condition, but no significant change in low risk CL condition.

	How much would you like to practise sport with…?		No significant change pre to post-test for social impact scores for low or high risk CL conditions.
	Observation of helping behaviour		No significant change pre to post- test for social preference or social impact of mainstream pupils in either the intervention or control conditions.
Bertucci, Conte, Johnson and Johnson (2010)	Classroom Life Instrument (Johnson, Johnson, & Anderson, 1983)	Perceptions of peer personal and academic support	Pupils who had worked in CL pairs and CL fours perceived greater peer academic support than pupils who had been in the individualistic learning condition. No significant differences between CL pairs and CL fours.
			No significant differences between CL and individualistic learning for peer personal support.
Bertucci, Johnson,	Classroom Life Instrument (Johnson, Johnson, & Anderson, 1983)	Perceptions of peer personal	Both CL conditions had greater perceptions of academic peer support than individualistic learning

Johnson and Conte (2011)		and academic support	CL resource and goal interdependence had greater perceptions of peer support than CL task and goal interdependence or individualistic learning
Bertucci, Johnson, Johnson and Conte, 2012)	Classroom Life Instrument (Johnson, Johnson, & Anderson, 1983)	Perceptions of peer personal and academic support	No significant differences between intervention groups (with / without group processing) for peer personal or peer academic support.
Jacobs et al. (1996)	A sociometric instrument. No further details provided	Friendship	Significantly greater gains in CL condition than control in Grade 4. No significant difference in gains between the two conditions in Grades 3 or 5.
Jacques, Wilton and Townsend (1998)	Sociometric: “name three children in your class you would like to invite to your birthday party”	Peer acceptance (of focus pupils with	Significant increases in social acceptance of the focus pupils by peers in CL classes but not in control classes This was maintained at 5 week follow up.

		‘mild intellectual disability’)	CL was as effective for focus pupils who had previously been in the mainstream classes as it was for focus pupils who had always been educated in special needs classes.
LaFont, Proeres and Vallet (2007)	Sociometric - number of attractions and rejections calculated “could you tell me which classmates you prefer / not prefer to play games with?” Sociogram (qualitative)	Interpersonal relationships (attraction, rejection)	No significant change from pre to post-test in number of attractions or rejections for intervention or control groups. Sociograms indicated increased attractions for 2 out of 3 experimental teams, and decrease for 2 out of 3 control teams. The least successful experimental team decreased, and the most successful control team increased. Least skilful players in the control group remained or became isolated by post-test.

Oortwijn et al. (2008)	Popularity: pupils asked to rate their team members as perceived by the whole class on the behavioural characteristic 'is well liked by everyone'. Perceived non-cooperativeness: at the end of each session, pupils recorded the names of team members whom they felt did not implement the CL rules. Totals for each nominated pupil were calculated.	Popularity Perceived non-cooperativeness	Following CL, there was a significant increase in ratings of popularity and significant decrease in perceived non-cooperativeness (pre to post). Significant increases in popularity and significant decreases in perceived non-cooperativeness for ethnically heterogenous teams, but not for ethnically homogeneous.
Piercy, Wilton and Townsend (2002)	Peer acceptance: 5 questions e.g. would you want to work with this child in class? using photos to indicate responses.	Peer acceptance Popularity	Significant increase pre to post-test for peer acceptance and popularity ratings for focus pupils with 'moderate to severe intellectual disability' by their peers in the CL

			condition, but no significant change in control or comparison condition.
	Popularity: from the photos, pupils chose 5 of their classmates they would most like to invite to a party at their home.		
Putnam et al. (1996)	Peer acceptance using sociometric method (Asher and Taylor, 1981) where mainstream pupils had to rate each classmate on a 5-point scale according to how much they would like to work with that person.	Peer acceptance (work)	Significant increase pre to post-test for peer acceptance in CL condition, but no significant change in control condition. This was true for mainstream pupils rating fellow mainstream pupils as well as when rating special education pupils. Peer acceptance in control conditions appeared very stable across the study period. This was true for mainstream pupils rating fellow mainstream pupils as well as when rating special education pupils.

Topping et al. (2011)	Sociometric: pupils indicated all pupils in their class according to three questions: like working with in science like spending time with at break time like seeing out of school	Friendships	No statistically significant differences found from pre to post-test for the intervention group regarding number of pupils like working with in science, like spending time with at break, and like spending time with out of school. At post-test, like working with was significantly greater in the control group than intervention At post-test, break time and out of school was significantly greater for the intervention than control group *CL appeared to enhance friendships beyond the work group but not in the work group.* Only two intervention schools had positive increases. Any positive findings at an individual class level almost equally balanced out by negative findings.
Vasileiadou (2009)	Sociometric	Peer acceptance	9 out of 17 children received an increased number of votes from peers

	Which children would you like to work / play / eat with? (by individual interviews and observations)	Of the 4 children identified as rejected, 3 received more votes from peers post intervention. The 4th child had no change in votes, but observations indicated increased playground interactions. The 3 children identified as high status initially, decreased but remained above the mean.

Appendix E Criteria used when judging the weight of evidence

	Low	Medium	High
WoE A ***Methodological Quality*** ▪ Aims / hypotheses / objectives clearly described ▪ Outcomes to be measured clearly described ▪ Participant characteristics clearly described (number, gender, age etc.) ▪ Random assignment to condition ▪ At least 20 participants in each condition ▪ Method / procedure clearly described ▪ Clear description of the intervention including duration, facilitator, approach taken to CL (e.g. Johnson and Johnson, Kagan etc.) ▪ Intervention discrete (no contamination or co intervention) ▪ Intervention fidelity described ▪ Control group +/or comparison group ▪ Established approaches to measuring outcomes ▪ Reliability and/or validity of relevant measures described ▪ Main findings clearly described ▪ Appropriate statistical methods used to analyse data ▪ Exact p values reported ▪ Limitations discussed	Fulfils 0-8 of the criteria	Fulfils 9-12 of the criteria	Fulfils 13-16 of the criteria

WoE B ***Methodological Appropriateness*** ▪ RCT or quasi-experimental design ▪ Random assignment (including group / block) ▪ Participants appropriate (age, setting etc.) ▪ Control or comparison group ▪ No additional or alternative treatments for experimental group ▪ Measures used relate to review question ▪ Pre and post-test data collected ▪ Data analysis is appropriate to the design ▪ Social skills training present ▪ Teacher selected heterogeneous groups	Fulfils 0-5 of the criteria	Fulfils 6-7 of the criteria	Fulfils 8-10 of the criteria
WoE C ***Relevance of Evidence*** ▪ Study based in the UK ▪ CL used as an intervention with whole class ▪ Delivered by class teachers ▪ Participants mainly in KS2 ▪ Participants are in a mainstream setting ▪ Participants are in their usual classes ▪ CL used within an academic lesson ▪ Study evaluates impact of CL on peer acceptance and/or class cohesiveness ▪ A treatment / teaching as usual control group	Fulfils 0-4 of the criteria	Fulfils 5-6 of the criteria	Fulfils 7-9 of the criteria

WoE D for each paper was calculated as the average of A, B, and C, or the most frequently occurring rating if an average was not possible.

Appendix F Weight of evidence analyses for the studies included in the systematic review

WoE A	Andre, Deneuve and Louvet (2011)	Andre, Louvet and Deneuve (2013)	Bertucci, Conte, Johnson and Johnson (2010)	Bertucci, Johnson, Johnson and Conte (2011)	Bertucci, Johnson, Johnson and Conte (2012)	Jacobs et al. (1996)	Jacques, Wilton and Townsend (1998)	LaFont, Proeres and Vallet (2007)	Oortwijn et al. (2008)	Piercy, Wilton and Townsend (2002)	Putnam et al. (1996)	Topping et al. (2011)	Vasileiadou (2009)
Aims / hypotheses / objectives clearly described	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Outcomes to be measured clearly described	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Participant characteristics clearly described (number, gender etc.)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Random assignment of participants to condition	Y	N	Y	Y	Y	Y	Y	Y	N	Y	N	N	N
At least 20 participants in each condition	Y	Y	N	N	Y	Y	N	N	Y	N	Y	Y	N
Method / procedure clearly described	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Clear description of the intervention e.g. duration, facilitator, approach taken to CL	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
Intervention discrete (no contamination or co intervention)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Intervention fidelity check(s) conducted	Y	N	Y	Y	N	N	N	N	Y	N	Y	Y	N
Control +/or comparison group	Y	N	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N
Established approaches to measuring outcomes	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
Reliability and/or validity of relevant measures discussed	Y	Y	Y	Y	Y	N	Y	N	Y	N	N	Y	N

Main findings clearly described	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Appropriate statistical methods used to analyse data	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Exact p values reported	N	N	Y	Y	Y	Y	N	N	N	Y	Y	Y	N
Limitations discussed	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Total score	15	12	15	15	14	13	13	11	13	13	15	15	8
WoE A rating	High	Med	High	High	High	High	High	Med	High	High	High	High	Low

WoE B	Andre, Deneuve and Louvet (2011)	Andre, Louvet and Deneuve (2013)	Bertucci, Conte, Johnson and Johnson (2010)	Bertucci, Johnson, Johnson and Conte (2011)	Bertucci, Johnson, Johnson and Conte (2012)	Jacobs et al. (1996)	Jacques, Wilton and Townsend (1998)	LaFont, Proeres and Vallet (2007)	Oortwijn et al. (2008)	Piercy, Wilton and Townsend (2002)	Putnam et al. (1996)	Topping et al. (2011)	Vasileiadou (2009)
RCT or quasi-experimental design	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Random assignment of participants to condition	Y	N	Y	Y	Y	Y	Y	Y	N	Y	N	N	N
Participants appropriate (age, setting etc.)	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N
Control or comparison group	Y	N	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N

No additional or alternative treatments for experimental group	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Social skills training specified as present	N	N	N	N	N	N	N	N	Y	N	N	Y	Y
Teacher selected heterogeneous groups	N	N	N	N	N	Y	Y	Y	N	Y	N	Y	Y
Measures used relate to review question	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Pre and post-test data collected	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Data analysis is appropriate to the design	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N
Total score	8	6	8	8	7	9	9	9	7	8	7	9	5
WoE B rating	High	Med	High	High	Med	High	High	High	Med	High	Med	High	Low

WoE C	Andre, Deneuve and Louvet (2011)	Andre, Louvet and Deneuve (2013)	Bertucci, Conte, Johnson and Johnson (2010)	Bertucci, Johnson, Johnson and Conte (2011)	Bertucci, Johnson, Johnson and Conte (2012)	Jacobs et al. (1996)	Jacques, Wilton and Townsend (1998)	LaFont, Proeres and Vallet (2007)	Oortwijn et al. (2008)	Piercy, Wilton and Townsend (2002)	Putnam et al. (1996)	Topping et al. (2011)	Vasileiadou (2009)
Study based in the UK	N	N	N	N	N	N	N	N	N	N	N	Y	N
CL used as an intervention with whole class	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N
Delivered by teaching staff	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	N
Participants mainly in KS2	N	N	N	N	Y	Y	Y	Y	N	N	N	N	N
Participants are in a mainstream setting	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Participants are in their usual classes	N	N	N	N	N	Y	N	N	Y	N	Y	Y	N

CL used within an academic lesson	N	N	Y	Y	Y	Y	Y	N	Y	N	Y	Y	N
Study evaluates impact of CL on peer acceptance (or related construct)	Y	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y	Y
Study evaluates impact of CL on peer support (or related construct)	N	N	Y	Y	Y	N	N	N	N	N	N	N	N
A control group (No CL)	Y	N	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N
Total score	5	4	6	6	5	8	6	6	6	4	7	8	2
WoE C rating	Med	Low	Med	Med	Med	High	Med	Med	Med	Med	High	High	Low

Weight of Evidence ratings given to the studies included in the systematic review

Study	WoE A	WoE B	WoE C	WoE D
Andre, Deneuve and Louvet (2011)	High	High	Medium	High
Andre, Louvet and Deneuve (2013)	Medium	Medium	Low	Medium
Bertucci, Conte, Johnson and Johnson (2010)	High	High	Medium	High
Bertucci, Johnson, Johnson and Conte (2011)	High	High	Medium	High
Bertucci, Johnson, Johnson and Conte (2012)	High	Medium	Medium	Medium
Jacobs et al. (1996)	High	High	High	High
Jacques, Wilton and Townsend (1998)	High	High	Medium	High
LaFont, Proeres and Vallet (2007)	Medium	High	Medium	Medium
Oortwijn et al. (2008)	High	Medium	Medium	Medium
Piercy, Wilton and Townsend (2002)	High	High	Medium	High
Putnam et al. (1996)	High	Medium	High	High

Topping et al (2011)	High	High	High	High
Vasileiadou (2009)	Low	Low	Low	Low

Appendix G Headteacher initial letter

Please ask for: Helen Smith (Trainee EP)
Direct Line: XXXX
E-mail: lpxhcs@nottingham.ac.uk
Fax No: XXXX
Date: 1st March 2016

Dear

I am writing to you as a second year trainee educational psychologist at the University of Nottingham, on placement in the XXXX Educational Psychology Service.

As part of my training, I am carrying out research into the impact of Cooperative Learning approaches in primary schools. Cooperative Learning is a method of structuring group activities in the existing, academic curriculum with the aim of promoting learning and enhancing children's social interaction skills.

I am looking for primary schools with **two or more class teachers in Year 4 and/or Year 5 who would like learn more about Cooperative Learning approaches** and would be willing to incorporate the techniques into their classroom in the Summer term. These approaches work within the curriculum and can be incorporated into existing lesson plans.

I have attached a question and answer information sheet giving more details about Cooperative Learning.

What are the benefits of taking part in the study?

For pupils, potential increases in:
- Academic Achievement
- Social Skills
- Self-esteem
- Social inclusion / Peer acceptance

For teachers and schools:
- Development of staff skills – **FREE TRAINING and professional development** for the class teachers involved in the study
- Ongoing support and monitoring during the research period, including visits to school
- Impact upon future practice
- A chance for your school to be involved in research

To be involved in this exciting opportunity, your school needs to have teachers:

- who are teaching in Years 4 or 5
- who are available for a training session in the second part of the Spring term
- who are willing to commit to implementing Cooperative Learning techniques in their classes during the Summer term
- who are happy for me to complete brief questionnaires with their pupils before and after the intervention
- who have not previously had training on or implemented Cooperative Learning approaches

For research purposes, it is important that there is at least one parallel year group class. This class teacher needs to be willing for their class to complete some brief questionnaires at the start and end of the research period and for me to visit to carry out brief observations of group work. These teachers will be offered full training in Cooperative Learning at the end of the research period.

Any identifying factors such as the names of pupils, staff and your school will be removed from the research report so that your involvement is anonymous. Confidentiality will be respected at all times.

Following the research, I would hope to discuss the findings with you; offering you a final report and an opportunity for further discussion. I hope that this would be a step towards building staff capacity and knowledge regarding Cooperative Learning approaches in your school.

I would be grateful for the opportunity to meet with you, or relevant person(s) in school to discuss the project and answer any questions you may have. I will contact you in the next week to discuss this opportunity further; however if you have any queries in the meantime, please do not hesitate to contact me.

Thank you for your time

Yours faithfully

Helen Smith
Trainee Educational Psychologist
XXXX Psychology Service
lpxhcs@nottingham.ac.uk

What is Cooperative Learning?

Cooperative Learning is a way of structuring elements of a lesson in order to encourage social interactions, whilst providing group members with the opportunity to practice their social skills.

'small groups of learners working together as a team to solve a problem, complete a task or accomplish a common goal' (Artz and Newman, 1990, p.448)

An example of Cooperative Learning is 'think pair share', where children are asked to discuss their answers or ideas with a partner, and then they both share their ideas with another pair of children.

Cooperative Learning approaches provide a structure for group activities which aim to:

- encourage positive interdependence – group members need each other to complete the task
- promote individual accountability – everybody has a stake in the final outcome, no freeloading!
- ensure equal participation – tasks are structured so that everyone can make an equal contribution
- actively teach relevant small group social skills through the academic curriculum

Cooperative Learning approaches can be easily fitted into the existing curriculum and existing lesson plans.

What does the research show?

Academic skills:

- increased achievement across subjects, age groups and ability levels
- enhanced problem solving
- increased time spent on task

Social and emotional development:

- enhanced self esteem
- increased motivation
- greater peer acceptance and pupil friendships
- increased acceptance of diversity
- higher behaviour ratings from staff
- enhanced liking of school

Source: Kagan and Kagan (2009)

Appendix H Information sheet for class teachers

School of Psychology
Information Sheet

*Title of Project***:** Cooperative Learning

Ethics Approval Number: 777

Researcher: Helen Smith (lpxhcs@nottingham.ac.uk)

University Supervisor: Nathan Lambert (nathan.lambert@nottingham.ac.uk)

Placement Supervisor: XXXX

This is an invitation to take part in a research project on Cooperative Learning.

Before you decide if you wish to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

If you are selected to participate as part of the **intervention group**, you will be expected to:

- attend a ½ day training session in the Spring term looking at social skills and Cooperative Learning structures
- use Cooperative Learning structures as part of a whole class lesson at least four times a week during the intervention period
- carry out the intervention over an 11 week period in the Summer term
- be part of ongoing monitoring and support visits by the researcher
- allow the researcher to collect measures using brief questionnaires with pupils before and after the intervention period,
- complete a brief questionnaire about your usual teaching practice at two points
- complete a brief, daily recording sheet to log Cooperative Learning sessions and pupil absence

If you are selected to participate as part of a **wait list condition**, you will be asked to:

- allow the researcher to collect measures using brief questionnaires with pupils before and after the intervention period
- complete a brief questionnaire about your usual teaching practice at two points
- be part of ongoing monitoring visits by the researcher

Participation in this study is totally voluntary and you are under no obligation to take part. You are free to withdraw at any point before or during the study. **All data collected will be kept confidential and used for research purposes only.** It will be stored in compliance with the Data Protection Act.

If you have any questions or concerns, please don't hesitate to ask now. I can also be contacted after your participation at the above address.

If you have any complaints about the study, please contact:
Stephen Jackson (Chair of Ethics Committee)
stephen.jackson@nottingham.ac.uk

Appendix I Consent form for class teachers

School of Psychology

Consent Form

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Title of Project: Cooperative Learning
Ethics Approval Number: 777
Researcher: Helen Smith (lpxhcs@nottingham.ac.uk)
University Supervisor: Nathan Lambert (nathan.lambert@nottingham.ac.uk)
Placement Supervisor: XXXX

The participant should answer these questions independently:

- Have you read and understood the Information Sheet? YES/NO
- Have you had the opportunity to ask questions about the study? YES/NO
- Have all your questions been answered satisfactorily? YES/NO
- Do you understand that you are free to withdraw from the study? YES/NO (at any time and without giving a reason)
- Do you give permission for your data from this study to be shared with other researchers, provided that your anonymity is completely protected? YES/NO
- Information from the study will be written up as part of doctoral research, and may be published in the future. Do you give permission for your data to be used in this way? YES/NO
- Do you agree to take part in the study? YES/NO

"This study has been explained to me to my satisfaction, and I agree to take part. I understand that I am free to withdraw at any time."

Signature of the Participant: Date:

Name (in block capitals)

I have explained the study to the above participant and he/she has agreed to take part.

Signature of researcher: Date:

Appendix J Information letter and consent form for parents

Please ask for: Helen Smith (Trainee EP)
Direct Line: XXXX
E-mail: lpxhcs@nottingham.ac.uk
Fax No: XXXX
Date:

Dear parent / carer

Cooperative Learning Project

Several schools in XXXX are taking part in a Cooperative Learning Project.

Your child's school has agreed to be involved in the project.

Cooperative Learning is way in which class teachers can organise group work during lessons. It aims to increase children's achievement, increase their time spent on task, increase their motivation and improve their self-esteem.

As part of my doctoral training as an Educational Psychologist, I am looking at the impact of Cooperative Leaning on peer acceptance and perceptions of peer support in the schools who are taking part in the project.

I am writing to request consent for your child to complete a brief questionnaire pack in their classroom (on two occasions).

Questionnaire information from the schools involved in the project will be used to look at the ways in which Cooperative Learning has been useful. This will form part of the write up of the research project.

The questionnaires ask children about how much they like to play and work with other children in their class, and how much they feel children in their class are supportive and helpful.

Responses will not be shared with school staff or pupils.

All involvement will be anonymous. The name of your child and school will be removed from the questionnaires.

If you are happy for your child to complete the questionnaires in class, please sign the consent form below and return to your child's class teacher as soon as possible.

If you change your mind at a later date, you are able to withdraw without having to give a reason.

If you have any questions about the research or would like more information, please don't hesitate to contact me.

Yours faithfully

Helen Smith
Trainee Educational Psychologist
lpxhcs@nottingham.ac.uk

PARENTAL CONSENT LETTER **Cooperative Learning Project**

School name:

Class teacher:

I give permission for **_____ (child's name**) to complete a brief questionnaire pack in their classroom in the Spring and Summer terms so that the impact of the Cooperative Learning approaches can be evaluated.

Parent / carer name: _____

Parent / carer signature: _____

Date: _____

Please return this consent slip to your child's class teacher

Appendix K Extracts from the CL training materials

The Cooperative Learning Project

Teaching Social Skills

and

the Cooperative Learning Structures

Contents

The Cooperative Learning Structures:

Numbered Heads Together

1. Put the children into mixed ability and gender groups of 4 (high, medium high, medium low, low).
2. Number children 1 – 4
3. Ask a question or set a problem to solve.
4. Pause to give children some individual think time
5. The children then 'put their heads together' to discuss their ideas and come to a consensus about the answer
6. **Make it clear to the children that all should participate and that by the end of the task, everyone in the group must be able to answer the question.** "Make sure that you can all…." Because no one knows which number will be called, all team members must be prepared.

Call a number (1-4) and children with that number take it in turns to share their group's answer to the class

Pick a table *and* child number. The other number 2s listen carefully to the child's answer and then have the opportunity to clarify the answer, add further information, or contradict parts of their answer.

Repeat the process with different questions

Appendix L Cooperative Learning adherence checklist

Date:

School / Teacher:

Intervention / Control - CL structure used:

Lesson:

Number of children: ***Adapted from Craig (2010)***

Area	Check	Present	Comment
Group	Appropriate size (max 4)		
	Heterogeneous groups		
	Groups selected by ***teacher pupil random***		
Environmental	Close proximity (can all see others)		
	Sufficient space		
	Materials available		
Staff	Task clearly explained		
	Objectives explicitly defined		
	Staff assistance 'guide on the side'		
Task	Positive interdependence		
	Individual accountability		
	Equal participation		
	Small group skills described		
	Reflection on small group skills		
Child	Simultaneous interaction		
	Follow instructions accurately		
	Solve conflict independently		

Appendix M Cooperative Learning daily log/record sheet (Intervention classes)

Cooperative Learning recording sheet **Week number:** **School:** **Year group:** **Teacher:**

Day	Date	• Structure(s) used • Number of children working together	Time spent using CL structure / activity	Curriculum area	Children who were absent for the session	Additional comments	Staff initials
Mon							
Tues							
Weds							
Thurs							
Fri							

Appendix N Cooperative Learning audit

Please indicate how often you use the following Cooperative Learning structures.

If you have not heard of a structure, that is fine.

Please select '7' for those items.

Think Pair Share _____

Timed Think Pair Share _____

Think Pair Share Squared _____

Timed Think Pair Share Squared _____

Numbered Heads Together _____

Travelling Heads Together _____

Quiz Quiz Trade _____

Inside / Outside Circle _____

Stand up, hands up, pair up _____

Rally Robin (pairs) _____

Round Robin (group) _____

Simultaneous Round Robin (group) _____

1 = daily
2 = most days
3 = once a week
4 = once a fortnight
5 = once a month
6 = rarely
7 = never

Please list / describe any other Cooperative Learning structures which you use in your classrom:

-
-
-
-
-

Name: _____

Signature: _____

Date: _____

Appendix O Classroom Supportiveness Scale (CSS)

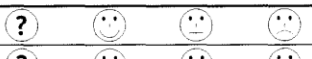
THE SENSE OF SCHOOL COMMUNITY SCALE PRIMARY

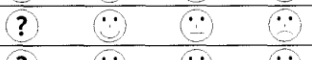
GL assessment
the measure of potential

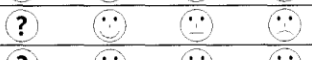
Name ..... Date .....

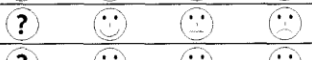
Age ..... Class ..... Please circle: Male / Female

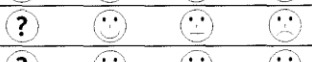
		Disagree a lot 1	Disagree a little 2	Neither agree nor disagree 3	Agree a little 4	Agree a lot 5
1.	Pupils in my class are willing to go out of their way to help someone.					
2.	My classmates care about my work just as much as their own.					
3.	My class is like a family.					
4.	The pupils in my class don't really care about each other.					
5.	A lot of pupils in my class like to put others down.					
6.	Pupils in my class help each other learn.					
7.	Pupils in my class help each other, even if they are not friends.					
8.	Pupils in my class don't get along together very well.					
9.	Pupils in my class just look out for themselves.					
10.	Pupils in my class are mean to each other.					
11.	When I'm having trouble with my schoolwork, at least one of my classmates will try to help.					
12.	Pupils in my class treat each other with respect.					
13.	Pupils in my class work together to solve problems.					
14.	When someone in my class does well, everyone in the class feels good.					
15.	When I'm having a problem, another pupil will help me.					
16.	Pupils at this school really care about each other.					
17.	Pupils at this school are willing to go out of their way to help someone.					
18.	Teachers and pupils treat each other with respect at this school.					

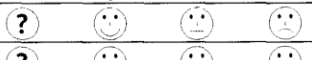


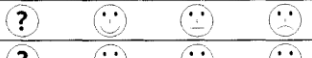


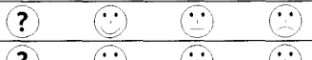


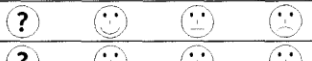


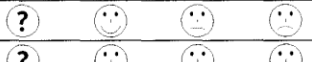


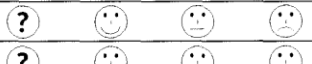


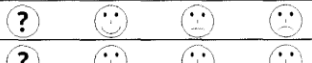


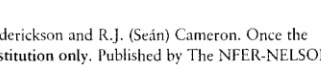


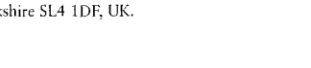












Appendix P Social Inclusion Survey (SIS)

NFER-NELSON
INFORMING YOUR DECISIONS

SOCIAL INCLUSION SURVEY (SIS)

How much do you like to .... with each person at school?

	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁
	?	🙂	😐	🙁

This measure is part of *Psychology in Education Portfolio* edited by Norah Frederickson and R.J. (Seán) Cameron. Once the invoice has been paid, it may be photocopied for use **within the purchasing institution only**. Published by The NFER-NELSON Publishing Company Ltd, Darville House, 2 Oxford Road East, Windsor, Berkshire SL4 1DF, UK.
Code 009000 7302

Appendix Q Information sheet for pupils

The Cooperative Learning Project

Information sheet for pupils

Your school and teachers have agreed to try out some new ways of organising your learning. Some of the teachers in your school will be trying these new ways in the Summer term.

I am going to try and work out whether these new ways of organising your learning are useful or not.

What does this mean for me?

I will come to your classroom on a few occasions. Once now, and again just before the Summer holidays.

I will ask you, and the other children in your class, to complete some short questionnaires.

What are the questionnaires about?

The questionnaires ask you about who you like to work and play with in your class, and how much you feel other children in your class are helpful and supportive.

Who will see my answers?

I will collect your questionnaires, and I will keep your answers private. I will not tell your teachers, parents or any children what answers you have put on your questionnaires. I will take your name away so that nobody knows they were your questionnaires.

This means that you can answer the questionnaires totally honestly.

I will make sure that I put your questionnaires in a safe place where only I can

see them.

What will happen next?

I will collect all of the information from everyone's questionnaires and look at it to see whether different ways of organising your learning are useful or not.

I will write a report which will tell people what I have found – but I will not name you, your friends, your teacher or your school.

Do I have to fill in the questionnaires?

Your parent / carers have said that you can fill the questionnaires in. But if you decide today that you do not want to, you can say no.

If you do not want to fill the questionnaires in, no one will be cross with you.

Even if you fill in the questionnaires, you can still change your mind later.

What do I do if I am happy to fill in the questionnaires?

In a minute, I am going to hand out the questionnaires. The first sheet asks you to write your name and the date. There is a box to tick to say that you are happy to fill in the questionnaires.

It will be very important to keep your answers private – so no peeking to see what other children put. There are no right and wrong answers on any of the questionnaires. I am just interested in what **you** think.

Before I talk you through the questionnaires, are there any questions?

Appendix R Pupil consent form 'willingness to participate'

Cooperative Learning Research

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Pupil Consent Form

'willingness to participate'

Name:

Date:

School:

Teacher:

I have read / had the information sheet for children read to me.

I understand that I am going to be given some questionnaires to help the researcher find out if different ways of organising my learning are useful.

I understand that I will be asked about how much I like to play and work with other children in my class, and how much I feel that children in my class are supportive and helpful.

I understand that I can choose whether to fill in the questionnaires or not, and that no one will be cross if I say no.

I understand that all my answers will be kept private and my questionnaire will be put in a safe place.

I have had a chance to ask questions.

☑ Please tick **one** of the boxes below:

I am happy to fill in the questionnaires ☐

I do not want to fill in the questionnaires ☐

Appendix S Ethics approval letter

UNITED KINGDOM • CHINA • MALAYSIA

School of Psychology
The University of Nottingham
University Park
Nottingham
NG7 2RD
tel: +44 (0)115 846 7403 or (0)115 951 4344

SJ/wb
Ref: 777

Monday, 18 January 2016

Dear Helen Smith & Nathan Lambert,

Ethics Committee Review

Thank you for submitting an account of your proposed research 'Cooperative Learning'.
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;

Reviewer 1:
1. In the access & recruitment letter for parents (pilot), please include a bit more information on the nature of Cooperative Learning and the questionnaire administered in class.

2. In the access & recruitment letter for parents, I would recommend including a brief description of what the control group entails (i.e., usual teaching) and what children in the comparison group will receive (e.g., increased shared learning experience). Providing a brief example of Cooperative Learning approaches would also be helpful.

Reviewer 2:

1. Official letters to schools or parents should not be written in Comic Sans font

2. Debrief forms should have contact details for the Researcher. It is not clear who they are for, but one presumes it is the teachers who delivered the intervention and the Head Teachers.

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 T Letters of thanks following the completion of the study

University of Nottingham
School of Psychology
Information Sheet for Parents/Carers
Research Project on Cooperative Learning
Ethics Approval Number: 777
Researcher: Helen Smith (lpxhcs@nottingham.ac.uk)
University Supervisor: Nathan Lambert (nathan.lambert@nottingham.ac.uk)
Placement Supervisor: XXXX

Dear Parents / Carers

You may remember that in early 2016, you gave consent for your child to complete some questionnaires as part of a research project that was taking place regarding the impact of Cooperative Learning. This project was carried out in four XXXX schools.

I am writing to inform you of the findings of the research, but would firstly like to take this opportunity to thank you for returning the consent slip agreeing for your child to complete the questionnaires.

The research investigated the impact of a whole class Cooperative Learning approach on children's peer acceptance and perceptions of peer support. At the start and end of the project, children completed questionnaires about who they liked to work and play with, and how much they felt that children in their class were supportive and helpful.

When the results from the four schools were combined, the project found positive impacts for the children who were in Cooperative Learning classes. By the end of the intervention, the children who were in Cooperative Learning classes had significant increases in

- their perceptions of peer support

- how much they like to work with pupils in their class of the same and different gender
- how much they like to play with pupils in their class of the different gender

There was a trend that children who were in Cooperative Learning classes liked to play with pupils in their class of the same gender more by the end of the intervention, but this was not statistically significantly different to any changes which occurred in the control classes.

I hope that this letter has provided some interesting feedback for you regarding the project. If you have any further questions, or would like a more detailed account of the subtleties of the findings, please do not hesitate to contact me, or speak to a member of staff in school.

Thank you once again.

Your faithfully

Helen Smith
Trainee Educational Psychologist
XXXX Psychology Service
lpxhcs@nottingham.ac.uk

University of Nottingham
School of Psychology

Information Sheet for Schools

Research Project on Cooperative Learning

Ethics Approval Number: 777

Researcher: Helen Smith (lpxhcs@nottingham.ac.uk)
University Supervisor: Nathan Lambert (nathan.lambert@nottingham.ac.uk)
Placement Supervisor: XXXX

Dear Staff

You will be aware that I was working in your school during the Spring and Summer terms 2016, in order to set up and evaluate a research project regarding Cooperative Learning.

I am writing to inform you of the findings of the research, but firstly would like to take this opportunity to thank you for agreeing for your school to be part of the project. I would also like to give particular thanks to the class teacher who worked hard in implementing Cooperative Learning approaches in the Summer term 2016.

The research investigated the effectiveness of a whole class Cooperative Learning intervention on children's peer acceptance and perceptions of peer support. At the start and end of the project, children completed questionnaires about who they liked to work and play with, and how much they felt that children in their class were supportive and helpful.

Four schools across XXXX took part in the research project.

When the results from the four schools were combined, the project found positive impacts for the children who were in Cooperative Learning classes. By the end of the intervention, the children who were in Cooperative Learning classes had significant increases in:

- their perceptions of peer support
- how much they like to work with pupils in their class of the same and opposite gender
- how much they like to play with pupils in their class of the opposite gender

These positive changes were statistically significantly different to any changes which occurred in the control classes, where teachers did not use Cooperative Learning.

There was a trend that children who were in Cooperative Learning classes liked to play with pupils in their class of the same gender more by the end of the intervention, but this was not statistically significantly different to any changes which occurred in the control classes.

I hope that this letter has provided some interesting feedback for you regarding the project, and that you will consider continuing to use Cooperative Learning techniques in the future. If you have any further questions or would like a more detailed account of the subtleties of the findings, please do not hesitate to contact me.

Thank you once again for your involvement in the project.

Your faithfully

Helen Smith
Trainee Educational Psychologist
XXXX Psychology Service
lpxhcs@nottingham.ac.uk

University of Nottingham

School of Psychology

Information Sheet for Pupils

Research Project on Cooperative Learning

Ethics Approval Number: 777

Researcher: Helen Smith (lpxhcs@nottingham.ac.uk)

University Supervisor: Nathan Lambert (nathan.lambert@nottingham.ac.uk)

Placement Supervisor: XXXX

Dear Pupil

You may remember that I came in to school last year, and you filled out some questionnaires about who you liked to work and play with, and how much you felt that children in your class were supportive and helpful. This was all part of the Cooperative Learning project.

What is this letter about?

I have now put together all the information from all the classes and all the schools, and am writing to let you know what I found out.

I would also like to say a big thank you to you for taking part.

What did the project find out?

Some really good things happened for the children who were in Cooperative Learning classes. When all the questionnaires from all of the schools were put together, the children who were in Cooperative Learning classes had significant increases in:

- how much they thought children in their class were helpful

and supportive

- how much they liked to work with pupils in their class of the same and different gender
- how much they liked to play with pupils in their class of the different gender

These positive changes seemed to happen for a range of different children in your classes.

What now?

I know that lots of you were keen to find out the results of the project, so I hope that this letter has given you some interesting information. I now hope that I will be able to help other teachers to use Cooperative Learning in their classrooms, and maybe help other children in other schools.

What do I do if I have more questions?

If you have any more questions, you can speak to one of your teachers in school. If they cannot answer your question, they will be able to get in touch with me.

Thank you once again.

Your faithfully

Helen Smith
Trainee Educational Psychologist
XXXX Psychology Service
lpxhcs@nottingham.ac.uk

Appendix U CSS data for individual schools

Table 6.1 Perceptions of peer support: Appended CSS data for schools A to D (means and change scores)

School	CSS Perceptions of Peer Support		*N*	Mean	*SD*	Change score	*SD*
A	Intervention	Pre Test	23	51.96	9.65	3.04	7.97
		Post Test	23	55.00	9.18		
	Control	Pre Test	27	46.30	8.71	1.00	8.26
		Post Test	27	47.30	9.80		
B	Intervention	Pre Test	24	46.67	7.07	1.29	7.17
		Post Test	24	47.96	9.22		
	Control	Pre Test	19	52.95	10.62	0.47	10.78
		Post Test	19	53.42	8.15		
C	Intervention	Pre Test	13	43.23	7.53	4.00	6.26
		Post Test	13	47.23	9.33		
	Control	Pre Test	19	44.42	8.76	-3.63	7.62
		Post Test	19	40.78	9.77		
D	Intervention	Pre Test	23	47.35	11.75	1.04	11.22
		Post Test	23	48.39	12.02		
	Control	Pre Test	22	47.05	6.95	-4.45	9.10
		Post Test	22	42.59	7.72		

Appendix V SIS WSG data for individual schools

Table 6.2 Peer acceptance: Appended WSG data for schools A and D (means and change scores)

School	**SIS WSG**		*N*	Mean	*SD*	Change Score	*SD*
A	Intervention	Pre Test	27	52.21	26.80	8.15	23.95
		Post Test	27	60.36	29.19		
	Control	Pre Test	28	37.34	44.70	-8.28	24.12
		Post Test	28	29.05	52.75		
D	Intervention	Pre Test	29	42.98	31.45	0.24	25.06
		Post Test	29	43.23	39.95		
	Control	Pre Test	28	64.06	27.71	-11.50	26.73
		Post Test	28	52.61	41.13		

Table 6.3 Peer acceptance: Appended WSG data for schools B and C (medians)

School	**SIS WSG**		*N*	Median	Range	Mean Rank
B	Intervention	Pre Test	25	54.55	118.18	25.52
		Post Test	25	54.55	172.73	23.80
	Control	Pre Test	26	68.75	138.89	26.46
		Post Test	26	66.67	177.78	28.12
C	Intervention	Pre Test	20	37.50	150.00	19.52
		Post Test	20	66.67	137.50	23.50
	Control	Pre Test	20	50.00	155.74	21.48
		Post Test	20	50.00	160.00	17.50

Appendix W SIS WDG data for individual schools

Table 6.4 Peer acceptance: Appended WDG data for schools A to D (means and change scores)

School	**SIS WDG**		*N*	Mean	*SD*	Change Score	*SD*
A	Intervention	Pre Test	27	16.55	29.42	4.72	21.48
		Post Test	27	21.30	29.74		
	Control	Pre Test	28	6.04	39.76	-2.38	26.88
		Post Test	28	3.67	47.96		
B	Intervention	Pre Test	25	-2.67	46.88	-5.00	21.65
		Post Test	25	-7.67	40.39		
	Control	Pre Test	26	14.62	43.39	-12.01	25.38
		Post Test	26	2.61	41.34		
C	Intervention	Pre Test	20	-8.47	47.49	35.42	36.10
		Post Test	20	26.95	34.65		
	Control	Pre Test	20	-7.50	35.43	-28.35	25.65
		Post Test	20	-35.85	24.86		
D	Intervention	Pre Test	29	-0.43	37.60	2.52	28.61

		Post Test	29	2.04	44.67		
	Control	Pre Test	28	25.06	25.28	-15.82	26.87
		Post Test	28	9.25	25.44		

Appendix X SIS PSG data for individual schools

Table 6.5 Peer acceptance: Appended PSG data for school A (means and change scores)

School	SIS PSG		*N*	Mean	*SD*	Change Score	*SD*
A	Intervention	Pre Test	27	55.99	27.99	-0.47	20.82
		Post Test	27	55.52	30.16		
	Control	Pre Test	28	26.42	40.02	0.34	25.87
		Post Test	28	26.76	54.20		

Table 6.6 Peer acceptance: Appended PSG data for schools B to D (medians)

School	SIS PSG		*N*	Median	Range	Mean Rank
B	Intervention	Pre Test	25	54.55	145.45	26.00
		Post Test	25	36.36	181.82	22.64
	Control	Pre Test	26	64.59	158.89	26.00
		Post Test	26	62.50	177.78	29.23
C	Intervention	Pre Test	20	33.33	125.00	17.60
		Post Test	20	64.59	150.00	21.18
	Control	Pre Test	20	42.86	160.00	23.40
		Post Test	20	60.00	150.00	19.82
D	Intervention	Pre Test	29	44.44	131	23.43
		Post Test	29	58.33	154	26.95
	Control	Pre Test	28	69.23	154	34.77
		Post Test	28	70.33	169	31.12

Appendix Y SIS PDG data for individual schools

Table 6.7 Peer acceptance: Appended PDG data for school A (medians)

SIS PDG		Condition	*N*	Median	Range	Mean Rank
School A	Intervention	Pre Test	27	-16.67	86.36	27.09
		Post Test	27	-8.33	105.55	28.85
	Control	Pre Test	28	-7.69	117.03	28.88
		Post Test	28	-7.42	128.57	27.18

Table 6.8 Peer acceptance: Appended PDG data for schools B, C and D (means and change scores)

SIS PDG		Condition	*N*	Mean	*SD*	Change Score	*SD*
School B	Intervention	Pre Test	25	-26.33	43.55	1.00	21.29
		Post Test	25	-25.33	39.78		
	Control	Pre Test	26	7.65	38.81	-6.20	25.80
		Post Test	26	1.45	45.48		
School C	Intervention	Pre Test	20	-25.97	47.10	39.58	35.99
		Post Test	20	13.61	36.68		
	Control	Pre Test	20	-48.35	27.21	-20.51	26.09
		Post Test	20	-68.86	18.55		
School D	Intervention	Pre Test	29	-23.66	37.90	22.10	39.13
		Post Test	29	-1.46	43.38		
	Control	Pre Test	28	21.37	25.16	3.29	24.77
		Post Test	28	24.55	23.06		