International Journal of Psychology and Educational Studies, 2018, 5 (3), 33-50

International Journal of Psychology and Educational Studies

www.ijpes.com

Interdependent Cross-Age Peer Tutoring in Mathematics

Mirjan Zeneli[1], Peter Tymms[2], David Bolden[2]

[1]*Universum College, Kosovë,.* [2]*Durham University, United Kingdom*

ARTICLE INFO

Article History:
Received 07.06.2018
Received in revised form 15.08.2018
Accepted 22.09.2018
Available online 10.10.2018

ABSTRACT

Peer tutoring is a form of structured peer learning technique. This study develops and tests a new form of peer tutoring technique, 'Interdependent Cross-Age Peer Tutoring' (ICAT). The method is informed by the 'what works literature' within peer tutoring and brings together crucial elements which have been shown to provide high effect sizes. Specifically, ICAT consists of an autonomous/informative structure, with students setting their goals in a cross age peer tutoring mathematics context. ICAT was implemented for six weeks in three different schools across England, with teachers concentrating on teaching their planed topics. School A (n=95) Year 8 students tutored Year 6, school B (n=65) Year 9 tutored Year 7, and school C (n=44) Year 10 students tutored year 8. Schools A and B adopt a pre/post-test quasi-experimental design and school C adopts a single group pre/post-test design. Research made instrument were applied to measure tutees performance gains. Classroom and paired observation were conducted for each school and the ICAT lesson materials for the six weeks were analysed to establish intervention fidelity. School A showed the highest ICAT implementation and effect size, 0.92, significant at (p<.001). However the findings should be interpreted with caution due to a weak research design.

Keywords:
S Social Interdependence Theory, Cross-age Peer Tutoring, Mathematics

1.Introduction

Meta-analyses have continually shown that cooperative group learning is an effective educational intervention (Johnson & Johnson 1989; Roseth, Johnson & Johnson 2008, Rohrbeck, Ginzburg-Block, Fantuzzo & Miller, 2003; Ginzburg-Block, Rohrbeck & Fantuzzo, 2006). Also, meta-analyses for students of ages 4-18 in peer tutoring have also shown to provide high effect sizes (Cohen, Kulik & Kulik, 1982; Leung, 2014; Zeneli).

Topping and Ehly (1998) provide elaborative peer tutoring typology. However, it can be broadly defined as a specific branch of peer learning in which students work on academic subjects in small groups, in which one or two students take the tutor or the tutee roles.

Abrahami et al., (1995) conclude that overall theories have tried to explain cooperative learning by concentrating on two elements: on students' motivation or on students' learning, and they provide an extensive 'organisational model' to understand and implement cooperative learning effectively. At the heart of the 'organisational model' is Social Interdependent Theory. The theory states that although having a peer learning strategy which concentrates on learning process is necessary to raise students' outcomes, - as Web (1989, 1992) and Topping and Ehly (2001) have argued, - motivating the peers to cooperate together and to learn is a precondition (Johnson, 1990; Johnson & Johnson, 1987; 1989; 2005; Johnson, Johnson, Holubec & Roy, 1984).

[1] Corresponding author's address: Rr. Imzot Nikprelaj p.n, Ulpianë, Prishtine
Telephone: +381 38 555 315
Fax: +377 44 144 062
e-mail:mirjan.zeneli@universum-ks.org
http://dx.doi.org/10.17220/ijpes.2018.03.004

Johnson, et al., (1984) identify four main cooperative skills levels that students need to be taught when working together; 1) forming, 2) functioning, 3) formulating and 4) fermenting skills. While the first and the second level deal with classroom arrangements, social and communication skills and motivation the last two levels concentrate mainly on learning skills such as seeking help, providing explanations, cognitive construction or knowledge reconstruction. At the heart of the theory are the ideas that formulating and fermenting skills are difficult to achieve without forming and functioning skills as well as applying social interdependent elements, such as reward interdependence, goal interdependence, task interdependence, role and interpersonal interdependence.

The following are some popular peer tutoring interventions:

1.1 Class-wide peer tutoring (CWPT)

There are at least 8 elements in a CWPT method; 1) Same-age and 2) similar-ability pairing, 3) reciprocity between peers, 4) structure, 5) method training, 6) group goal, 7) group contingency, and 8) scripts with instructions and answers (Arreaga-Mayer, Terry & Greenwood, 1999). Beyond these 8 elements, CWPT has incorporated elements such as 'praise' and different cognitive or meta-cognitive strategies. The next two methods, RPT and PALS derive from CWPT:

1.2 Reciprocal peer tutoring (RPT)

Reciprocal Peer Tutoring is similar to CWPT, its main difference lies in the pairing format and the form of goal and reward structure. Whereas in CWPT the pairs are assigned randomly and then allocated to one of two groups and compete against one another for a reward set by the teacher, in RPT the pairs are not randomly created, rather it is the pair who chooses the type of goal and reward from a list (Fantuzzo, King & Heller, 1992).

1.3 Paired-assisted learning strategies (PALS)

Paired-assisted learning strategies contain all the above CWPT characteristics, however it is different to CWPT in at least four characteristics: 1) Even though same age, the pairs are cross ability, 2) PALS incorporate a modelling characteristic, in which the higher ability student leads the way, 3) emphasis is placed on praising, 4) cognitive and meta-cognitive strategies of learning are applied (Fuchs, Fuchs, Karns, Hamlett, Katzarokt & Dudka, 1998).

1.4 Cross-age peer tutoring

In a cross-age peer tutoring the tutor and the tutee differ from one another in terms of age, and consequently in ability level, with the older peers possessing a more advanced knowledge of the materials. Also, due to its nature, cross-age tutoring is always fixed role, with the older student acquiring the role of the tutor (Sharpley & Sharpley, 1981). Beyond these three characteristics, age, ability and role, the method can have any combination of the remaining characteristics associated with peer tutoring.

Many studies and suggestions exist on simplistic cross age peer tutoring (Fitz-Gibbon, 2000; Thurston 2014; Tymms, Merrell, Thurston, Andor, Topping & Miller, 2011; Tymms & Merrell 2015)

1.5 Significance

Cross-age peer tutoring according to Fittz-Gibbon is superior to cross-ability same-age peer tutoring, as this method threatens the student who seeks help, whereas help from an older student is more accepted (Fitz Gibbon, 1992). One of the earliest studies to look into this is that by DePaulo et al., (1989), concluding that a) students tutoring same-age peers feel threatened by the task, b) cross-age tutors are also threatened if they are of similar ability, and that c), tutees can feel threatened to ask for help in a same-age context. Studies have shown that cross-age peer tutoring is supervisor than same-age peer tutoring in improving students' attitude (Miller, Topping & Thurston, 2011) and performance (Cohen, et al., 1982; Tymms, et al., 2011). Moreover, meta-analyses in peer tutoring have shown that peer tutoring informed by the social interdependent approach yield positive attainment results (Leung, 2014; Zeneli, Thusrton, Roseth, 2016). Specifically, studies with goal interdependence set by students provide higher effect size=.99, as opposed to those on which the teacher sets the goals ES=.30, as do studies with autonomous structures ES=.94, without autonomy ES=.30 (Rohrbeck, Ginsburg-Block, Fantuzzo & Miller, 2003). Also, meta-analysis in peer learning in general by Roseth, Johnson and Johnson (2008) have shown that positive peer relations predicts outcome performance on peer learning, R^2=.40.

Positive results for the impact of an interdependent cross-age peer tutoring (ICAT) on tutees' attitudes have been illustrated (Zeneli, Tymms & Bolden, 2016). However to date researchers have not looked at applying important social interdependent elements to cross-age peer tutoring to measure its impact on attainment results, specifically important elements such as social skills training, goal/reward interdependence and autonomous/informative structure interdependence (Zeneli & Tymms, 2015).

Aim and research questions

This paper aims to apply key social interdependent elements to cross age peer tutoring in mathematics. Specifically it tries to answer the following research questions:

A). Does cross age peer tutoring raise tutee's mathematics attainment when incorporating crucial social interdependence elements such as goal, interpersonal training and resource interdependence, academic/social guidance scripts and autonomous structure?

B) To what extent was the intervention implemented successfully across the schools?

3. Method:

3.1 Participants

There were in total 550 students from three schools across England that participated in this study; two in the North East of England and one in the South East. The schools in the North East involved Year 8 tutoring Year 6, (School A), and the Year 9 tutoring Year 7 (School B). The school in the South East involved Year 10 tutoring Year 8 (School C).

The schools for the project were chosen as follows. Emails reporting the opportunity for the project were sent to the school authorities of three different English county councils, North Tyneside, Leeds, and Medway, as well as the secondary schools working with the Centre for Evaluation and Monitoring (CEM) at Durham University. Over 70 schools were contacted. The response rate of schools who expressed the wish to participate in the project was 9 schools, 13% of the total contacted. Out of the 9 schools only 3 agreed to the terms and conditions in respect to design and time frame.

Tables one provides more information regarding the school characteristics and group gender context. Although Special Education Needs students participated in the project they were not included from the analysis in order to keep a homogeneous sample.

Table 1. School characteristics

	Mean Age For Each Year					School Age Range	Total Students	% Boys	% Girls	% SEN*	% ESL*	% Free School Meal	Average KS2 Point
Schools	6	7	8	9	10								
A	11.7		13.4			9 -13	478	50.6	49.4	5.2	7.9	26.6	28.8
B		12.3		14.2		11 – 16	415	53	47	11.3	0	54.5	28
C			13.2		15.4	11 – 19	1301	55.1	44.9	13.7	1.6	28.5	26.5
National Average							978	51.0	49.0	7.7	15.9	16.8	28.4

**SEN=special educational needs. ESL=English as second language*

For an expected effect size .55, which is the average effect size found in peer tutoring interventions in mathematics (Zeneli, 2015) significance level of .05 (one tailed T-test), and a power of .8, the total sample requirement for each school is 84, or 42 per each group. In order to account for attrition rates it was necessary to have just over 50 participants per each group, however for school B the numbers were smaller as there were not enough students.

3.2 Design

Initially all the schools agreed to participate on a quasi-experimental design intervention, however, school C decided to drop the control group and applied the intervention to the entire age group, without conducting any post-tests on the control classes. Table two presents the design and the data collected, and table three next page presents the gender composition in each group.

Regarding group characteristics in terms of free school meal (FSM), for school A the control group consisted of 8 FSM students, and 6 for the peer tutoring group. For school B, the control group contained 17 FSM students, and the peer tutoring group of 15 students. Finally, for school C there were 19 FSM students.

Table 2. Design and data collected

Schools	Length & control groups.	Performance	Attitude Data*	Observation	Lesson scripts
A	Six weeks, pre-post-test quasi experimental design.	Pre-post for all Year 6s.	Pre-post, for all Year 6s and Year 8s.	Two classes.	Two classes out of four.
B	Six weeks, pre/post-test quasi experimental design.	Pre-post for all Year 9s and Year 7.	Pre-post, for all Year 9s and Year 7s.	Two classes.	Yes.
C	Six weeks, pre-post-test single group design.	Pre-post for the Year 8s.	Pre-post, for peer tutoring Year 10s and Year 8s.	Two classes.	Yes.

*Attitude data are not reported in this paper

Table 3. Gender composition for each group by school and year

School A				*School B*				*School C*			
Year 6	Control	Girls	24	Year 7	Control	Girls	22	Peer Tutoring	Year 8	Girls	31
		Boys	34			Boys	17			Boys	43
	Peer Tutoring	Girls	22		Peer Tutoring	Girls	23		Year 10	Girls	36
		Boys	32			Boys	13			Boys	44
Year 8	Control	Girls	41	Year 9	Control	Girls	18				
		Boys	21			Boys	24				
	Peer Tutoring	Girls	25		Peer Tutoring	Girls	15				
		Boys	29			Boys	21				

3.3. Intervention Procedures

Organisation. Every school agreed to 35-40 minutes of cross age peer tutoring for six weeks, as well as 45 minutes of student training prior the intervention. The project started at the last term of the 2013 academic year. A timetable with all data-collection times was provided to each school. The short tests were administered and collected by the teachers together with all peer tutoring lesson scripts. The observations were conducted by the first author.

Materials/exercises. The topics of the materials were chosen by the schools in order to reflect their lesson plans, hence the control group also concentrated on the same topics. The exercises were created by the mathematics teachers from each school, together with the researchers. Many of the exercises were influenced by MathsLinks, 1, Year 7 Practice Book (Allan, 2008), for it provides a good illustration how to range mathematics exercises in different complexity levels, a crucial component in the interdependent cross-age peer tutoring (ICAT) framework.

Table four next page provides additional information on the exercises covered by each school:

Table 4. Lesson topics by schools

School A	School B	School C
1. Number patterns and sequence 2. Fractions 3. Understanding measures 4. Properties of shape 5. Data interpretation and representation 6. Written methods	1. Mean, median, range, mode 2. Data interpretation and representation 3. Factors, multiply, fractions 4. Sequences 5. Mental methods, multiply/divide 6. Equations	1. Measurements 2. Probability 3. Transformation 4. Enlargements 5. Area 6. Equations

Pair and Class set up. The pairings were conducted by the mathematics head teachers based on previous classroom individual assessments. The pairings of the students were conducted in the following way: The highest performer of the older age was paired with the highest performer of the younger age, and so on down the line on the same sex. This form of pairing on ability is consistent with previous cross age tutoring interventions (Fitz-Gibbon 2000; Topping, K. J., Campbell, Douglas & Smith, 2003; Tymms, et al., 2011), and has also been recommended by Thurston's (2014) peer tutoring manual supported by ESRC and Tymms and Merrell (2015). Also meta-analysis have illustrated that same sex grouping provide higher effect size (Rohrbeck, et al., 2003). Once the students were paired, they worked together for the duration of the project.

3.4. Training

Mathematics head teachers and in-house facilitators, teacher and students all received ICAT training. The training for the mathematics head teachers and facilitators concentrated on the following three areas:

Theory. Why and how peer tutoring works, and the literature that exists on peer tutoring.

Practice. Individual role-play on the ICAT framework.

How to train the teachers and the students to concentrate on academic and social interaction skills: Specifically how to give praise correctly and synchronise the tone of voice with the body language and the context overall. The head teachers and the facilitators were advised to carry out role play with the teachers and the students on both: the academic framework and interpersonal communication skills. In other words the training was conducted in a cascade top-down model.

Prior to the intervention the students received one full training session of 45 minutes, working on simple mathematics materials. Studies have found that training of the students is essential in peer tutoring (Harrison & Cohen, 1971; Barron & Foot, 1991).

The 45 minutes student training was divided as following:

The first 20 minutes of the training concentrated on interpersonal communication skills development. Specifically, the pairs were asked to do role playing: First they were required to sit far away from each other, interrupt and maintain a neutral face, - the pair then sat close to each other, smile, and be nice to one another, - and finally to discuss why the second option is better and the importance of synchronising body language with the tone of voice.

The remaining 25 minutes of the training, concentrated on the peer tutoring scripts, going through with their teacher what to do at each stage.

3.5. ICAT intervention framework

ICAT framework developed here consisted of four parts: As it can be seen in the ICAT framework, figures one, two and three,

The steps that the students took in order to complete the work were as following:

Part 1, Goal Setting. In the first part, Goal Setting, the students chose together in pairs a number threshold that they wished to reach, a threshold which become their performance goal.

Part 2, Practice-Test. In the second part, Practice-Test, the tutor prompted the tutee to answer a range of mathematics questions. This was the part in which new mathematics concepts were introduced to the tutee for the first time, ranging from very easy to very difficult. Maximum interaction was expected at this stage, as the tutee would struggle with new concepts, and the tutor was expected to provide help in different ways, first implicitly then explicitly.

Part 3, Connect. High interaction was also expected at the third stage, 'connect'. At this stage the tutee was prompted by the tutor to connect the new concepts to previous mathematics concepts and to real life events. This part ensured advanced cognitive and meta-cognitive engagement, and the tutor was also asked to provide help in different ways; again first implicitly then explicitly.

Part 4, Turn-Taking Test. The final part, contained less verbal interaction and more tutor written modelling, as the students were required to take turns to complete the exercises in order to determine whether they had achieved their self-set performance goals. By making the tutor engage in exercises in the end as well, it was expected that the tutor would have to take stages 1, 2, and 3 seriously, as he/she would need to have a good engagement with the mathematics concepts as well.

GOAL: Together, choose a number between 5–15. Write the number inside this circle: ◯
This is the number of points you both want to win. Together you will do a little test at the end. Each correctly completed exercise will earn you 4 points.

PRACTICE TEST: **Tutee's Role:**	**Tutor's role:**
2). Try solving: Find the missing numbers: 1. $8 + ? = 12$ 2. $3x = 18$ 3. $b - 3 = 12$ 4. $2x + 4 = 12$ 5. $3a - 4 = -10$ 6. $4x + 2 = 2x + 18$ 7. $6(a + 2) = 4(a + 4)$ *Clues:* ↓	1). Ask: "How can we think about these exercises and solve them?" Direct only with nods. If correct praise *kindly*: Well Done! Very Good! Excellent! Brilliant

3). *Ask:* Can you give me any clues? *For an answer:* ↓	4). Respond with *only* hints and clues. If correct, praise *kindly:* Well Done! Very Good! Excellent! Brilliant!
5). *Ask*: How would you answer?	6). Try giving *answers*. Once written down, praise *kindly:* Well Done! Very Good! Excellent! Brilliant

Figure 1. Parts one and two

CONNECT: **Tutee's Role:**	**Tutor's Role:**
2). Summarise ideas below: *Clues:*	1). Ask: "Can you think of other things you have done in maths like this? "Can you think when you might do this kind of maths in your everyday life?" Direct only by nods. If correct, praise *kindly*. Well Done! Very Good! Excellent! Brilliant
3). *Ask:* Can you give me any clues? *For an answer:*	4). Respond with only hints and clues. If correct, praise *kindly*. Well Done! Very Good! Excellent! Brilliant
5). *Ask:* How would you answer?	6). Try giving ANSWERS. If written down, praise *kindly*. Well Done! Very Good! Excellent! Brilliant! **If you are not sure ask the teacher.**
TURN TAKING TEST: 1) Ask the teacher for the test. 2) Do the exercises in the test by taking turns. 3). Tutor marks tutee's exercises and gives feedback. 4) Together, add up the points for every *correct* answer, and check if you have won. 5) High-five each other for the great effort. ☺	

Figure 2. *Parts three and four*

Solve the exercises below by taking turns:

1. $8 + ? = 21$
2. $3c = 21$
3. $b - 3 = 15$
4. $2x + 4 = 18$
5. $3a - 4 = -13$
6. $4x + 2 = 2x + 22$
7. $4(c + 4) = 6(c + 2)$

Figure 3. Part four

The intervention adopted here was different from other cross-age peer tutoring interventions when considered as a whole. The following are some of ICAT's strengths:

Academic goal interdependence. Unlike most *cross-age* peer tutoring interventions, which rely only on the role and social skills training as means of positive interdependence, the framework here also incorporates academic-goal interdependence.

Autonomous-Structural interdependence. This was achieved by combining guidance sheet, exercises and praising cards all in one. This is unique to the ICAT framework. An evident problem in peer tutoring is that tutors give the answer to the tutee too early (Harrison & Cohen, 1971), therefore the timing of when the answer be given is very crucial according to Topping (2001); similarly the timing of praising is also important (Johnson, 1990).

Advanced cognitive and metacognitive engagement. ICAT explicitly engages students in meta-cognitive discussion by providing a section entitled 'connect', in which the students are asked to engage into two different, yet similar, ways of thinking: First to connect what they just covered to previous Mathematics classes or subjects in order to provide some kind of *categorisation* (Kramarski & Mevarech 2003; 2004), and second to relate the topic to real life events, as in the case of "link it up" in Shared Maths peer tutoring (Tymms, et al., 2011). Both relating it back to current knowledge or to real life are emphasised and advised by Thurston (2014) peer tutoring manual.

It familiarises students with tests by providing a test-like peer tutoring environment. This was achieved by naming two out of four peer tutoring sections as; "Practice-Test" and "Turn-Taking-Test". Again this is similar to Fantuzzo, Polite and Grayson (1992) Class-wide Peer Tutoring or to RPT in general, however, with three major differences. Firstly, the exercises, for both parts, range hugely in the level of complexity, hence aiding the students to reach their true ZPD; this is important since in cross age peer tutoring the teachers very often do not know what level the tutors or tutees are at, teachers often find it hard to provide the students with the correct exercise levels. Secondly, the final test is a turn-taking test, hence the students carry on working together, rather than alone. They do so, however, in such a format that the tutee and the tutor both watch how their partner solves the exercises, so that there is room for peer modelling in the process.

3.6. Implementation indicators

The table below shows the attrition rate for each group and data analysis type for each school:

Observations: Observations were conducted at two levels, overall observations and student pair observations.

Overall observations: This concentrated at the classroom level, ensuring that basic elements were met, those were: Spacious classrooms, teachers knowledgeable in terms of ICAT elements, teachers helping students, desks organised appropriately, necessary materials were present, and same sex pairings.

Pair observations: The observation instruments, table five, have been influenced by various thinkers, items 1-3 by Argyle, (1976); Allen and Feldman, (1976), Johnson (1990) measuring the level of interdependence, while the remaining items have been influenced by Fitz-Gibbon (2000), Roscoe & Chi, (2007); Topping, Miller, Murray & Conlin (2011), measuring the level of cognitive and meta-cognitive engagement.

The structure of the observation is influenced by Topping, et al., (2011). Observations were conducted in the following way: Each pair of students was observed through five windows. For the Year 8 tutoring Year 6, and Year 9 tutoring Year 7 each window lasted approximately 30 seconds, 5 seconds to adjust, 15 seconds to observe and 10 seconds to record. For the school of Year 9 tutoring Year 7 the window was longer, as there were only 9 pairs per class. Therefore for the Year 9 tutoring Year 7 each window lasted around 42 seconds.

Lesson Materials: Lesson materials were collected for most of the 6 lessons, including the training lessons for school B and C.

The lesson materials concentrated on measuring different areas such as: first, the level of interdependence, by paying attention to the number of students setting goals, second, cognitive and meta-cognitive elements, by paying attention to the quantity of exercise attempted at each level, the quality of answers and the quantity of different feedback types.

Table 5. Implementation indicators by method

	Lesson Materials		Observations
1	Amount of peer tutoring lessons scripts including training.	**1**	Goal Interdependence
2	Set goals by the students	**2**	Tutor B-Language/T-Voice
3	Quantity of exercises attempted in the *practice test* section	**3**	Tutor Praises Correctly
4	Quantity of exercises attempted in the *connection* section	**4**	Tutor M/C Questions
5	Quantity of exercises attempted in the *turn-taking* test	**5**	Tutee Self Corrects
6	Quality of answers in the *practice test* section	**6**	Tutee Connects/Categorises
7	Quality of answers in the *connection* section, such as negative, broad or specific statements	**7**	Tutor Questions
8	Quality of answers in the *turn-taking* test	**8**	Tutee Answers
9	Total types of feedback	**9**	Tutee Questions
10	Feedback by ticks/crosses in the *practice test*	**10**	Tutor Explains
11	Feedback by ticks/crosses in the *turn-taking* test		
12	Checking if goal is achieved.		

3.7. Instruments

Since this was a short intervention, it was decided that in order to capture the effect of the intervention the instrument needed to be newly designed as oppose to a standardised test which would have been too broad to capture the impact of the intervention.

Also, since the students worked at different levels due to the difference in age across schools and different topics, three different instruments were created. Each instrument measured mathematics performance of the topics covered during the 6 weeks that took place in the experimental and the control groups. The following were the Cronbach's Alpha reliability levels presented in table 6:

Table 6. Cronbach's Alpha reliability coefficient for performance measurement

Schools	*Year*	*Cronbach's Alpha Based on Standardized Items*	*n of Items*	*n*
	Year 6 Pre-test	.70	6	98
A	Year 6 Post-test	.67	6	97
	Year 7 Pre-test	.66	5	64
B	Year 7 Post-test	.70	5	63
	Year 8 Pre-test	.36	6	49
C	Year 8 Post-test	.59	6	46

3.8. Analysis

Performance was analysed by using SPSS 20:

Analysis of Covariance (ANCOVA) was used for schools A and B, the pre/post-test quasi experimental designs. ANCOVA controls for pre-test score differences.

For school C, the single group design a dependent t-test was applied. SPSS does not conduct one sided tale t-tests, which was required when testing rather than exploring hypothesis; therefore in order to take account of this shortfall the p value was multiplied by 2 for the dependent t-test.

The Effect Sizes for the quasi-experimental designs was manually calculated by taking the ANCOVA coefficients and using Cohen's d technique with the square root of the MSerror as the denominator. Hedge's g pooled standard deviation was used to calculate the effect size for school C as the design differed from schools A and C.

3.9. Results

3.9.1 Attrition rate

Table 7 reports the attrition rates by data collection type:

Table 7. Attrition rate % for each data collection type by group and school for the tutees

Schools	Groups	Performance tests %*	Lesson materials %*
School A			
Year 6 students	Peer Tutoring	9 missing 9/54=17%	31 missing students 31/54=57%
	Control	9 missing 9/58=16%	
School B Year 7 students	Peer Tutoring	1 missing 1/36=3%	1 missing student 1/36=3%
	Control	13 missing 13/39=34%	
School C			
Year 8 students	Peer Tutoring	30 missing 30/74=41%	30 missing students 30/74=41%

**In order to arrive at the attrition percentage the number of missing students was divided by the total number of students within each group for each data collection type.*

In terms of performance data the highest attrition rates were for school B the control group, 34%, and school C with 41%. In terms of lesson materials, more than half of school A students, 57%, did not provide their names in the ICAT lessons therefore they did not enter the analysis; and 41% of students in school C did not do so.

3.9.2 Implementation fidelity

Observations

Overall school/classroom observations experience. For school A the observations of the set up overall corresponded to the planned intervention. The classes were spacious, and there was enough space between the pairs. Teachers directed the pace of the peer tutoring as they were trained to do, guiding the pairs through each peer tutoring part. The lessons scripts indicated that the schools made use of the training lesson. The pairing of the students was conducted as planned, i.e. same sex and the best performer of the older age was paired with the best performer of the younger age.

School B showed few problems: One of the classrooms was slightly small. Even though there were only nine pairs the students sat very close to one another. Although the teacher did try his best to guide the pairs through each peer tutoring part many of the students did not remain seated, hence causing noise. The pairing of the students was only partly conducted as planned, specifically the school did not entirely managed to secure same-sex pairing. Also, the school teachers were trying Information Communication Technology (ICT) and formative feedback teaching strategies in order to improve students' performance, therefore the context in this school was not normal.

School C the observations revealed that overall the classroom sizes and pair's seating were spacious. However, there were other issues: Firstly, this school had entirely split the boys from the girls so as to be sitting in different rooms, this was not an issue, however, for the boys' observation, there were three teachers in the room: the appointed teacher, the mathematics head teacher and the school's head teacher; showing extra effort in implementing peer tutoring. Later, however, it became evident that on that same day the school had been inspected by Ofsted, who are usually keen on peer tutoring interventions. Therefore, the schools' effort to manage peer tutoring effectively is very likely to have been an extra effort in light of the Ofsted inspection rather than a genuine attempt.

Table 8 next page provides an overall picture of the general school observations:

Table 8. General school observations

Observations topics	School A	School B	School C
Classrooms space	Enough space	One classroom too small	Enough space
Teachers knowledgeable with ICAT	Yes	Yes	Yes
Teachers aiding students	Yes	Yes	Yes
Desks organized appropriately	Yes	Mostly	Yes
Materials (pencils, rulers, etc.)	Yes	Yes	Yes
Same sex pairing	Yes	Mostly	Yes

Pair observations: Figure four shows the extent to which student pairs implemented ICAT according to programme specifications.
Overall the pair observations illustrate that the indicators least implemented were goal interdependence, tutor praises correctly and tutee connects/categorises. Also, indicator 'tutee self-correct' was removed from the analysis since this activity did not materialise within the pair. The indicators with the highest observed frequency were: tutor adopts a correct body language and tone of voice, tutor explains, tutor asks questions and tutee provides answers/explanations.

In terms of comparing the schools to one another, for most of the indicators no major differences were detected, with the exception of school A in which more goals and correct praising were observed; and in school C the tutors seemed to ask more questions compared to other schools.

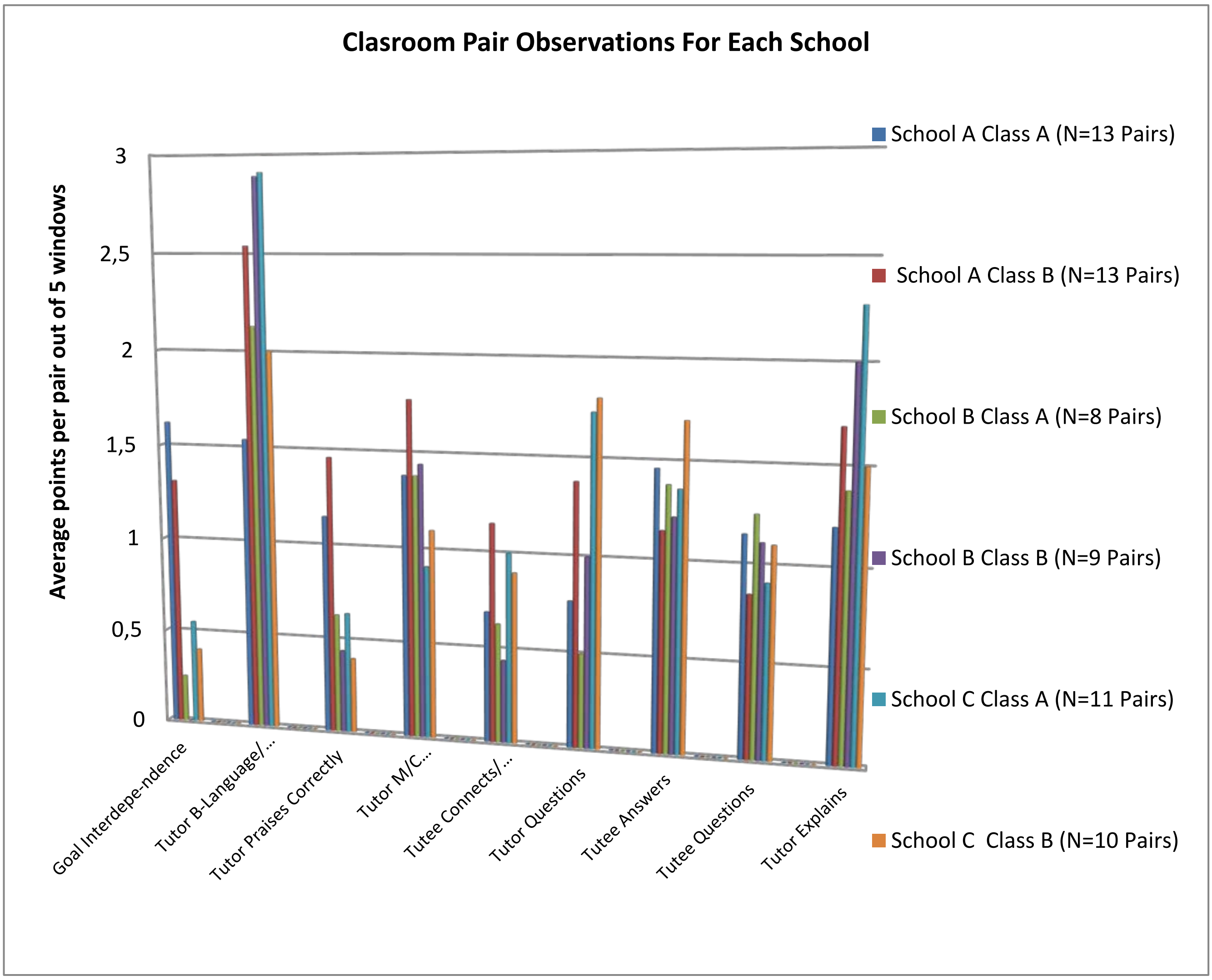

Figure 4. Classroom pair observation by school

Lesson materials: In terms of ICAT lesson materials, school A came the closest to implementing the programme according to its specifications, 85%. Schools B and C showed an implementation of 63.5% and 64.88% respectively. Table 9 provides the results.

Overall the element least implemented was 'lessons with student set goal', 58%; and the element implemented the most was 'lessons in which all exercises in the practice-test section were attempted', 79%.

Table 9. Percentage of implemented lessons according to programme specifications for the students with their names on the ICAT materials.

Lesson Materials	School A %	School B %	School C %	Average % per indicator
Total peer tutoring lessons attended with students names on.	149/161=93	152/245=62	206/308=67	74
Lessons with student set goals	145/149=97	60/152=39	80/206=39	58
Lessons in which all exercises in the practice-test section were attempted	116/149=78	120/152=79	164/206=80	79

Lessons with attempted exercises connect section	105/149=70	85/152=56	154/206=75	67
Lessons in which all exercises in the turn-taking section were attempted	121/149=81	106/152=70	120/206=58	70
Correct answers for practice test	779/957=81	632/986=64	1410/1759=80	75
Specific answers connect section	363/382=95	92/130=71	124/256=48	71
Correct answers turn-taking section	785/942=83	651/967=67	819/1132=72	74
Average per school	85	64	65	

3.10. Performance

Statistical assumptions: There were two statistical assumptions which were not met when conducting the ANCOVAS. The violated assumptions were found in school A, when comparing the mathematics performance of the Year 6 peer tutoring group to the Year 6 control group; those were normality assumption for the peer tutoring group on the pre-test and the homogeneity of variance between the groups. However, when the number of participants is over 30, as was the case here, assumption violations do not have a huge impact on the findings (Howell, 2010).

Pre-test differences: Table 10 suggests that there were significant differences for the pre-test data for school A. Hence to ensure that there was no regression to the mean, a condition in which a particular group falls within the extreme side of the sample mean in the pre-test and then equalises in the post test, the formula recommended by (Trochim, 2012) was applied: 100(1-r), where r is the correlation between the pre and post-test data within each group. The higher the correlation, the lower the percentage of the regression to the mean (Trochim, 2012).

All groups showed a 20-30% regression to the mean. However, since this phenomenon appeared for both groups, control and peer tutoring, the threat to inflating the effect size is extremely small, as regression to the mean for any variable has cancelled itself out, and therefore poses no major danger to the results.

Effect sizes. Table 11 provides details on the findings. The highest ES was that of year 6 tutees within school A as compared to the control group within that school, ES=.92, significant at ($p<.001$). This was followed by Year 8 tutees within school C, comparing the peer tutoring pre-tests scores to the post-tests, ES=.79, also significant at ($p<.001$). Finally the lowest ES was that of Year 7 tutees within school B as compared to the control group within that school, ES=.22, non-significant.

Tables 10 and 11 next page provide additional data:

Table 10. Pre-test score differences

Student Performance	*Pre-test mean peer tutoring*	*Std.*	*n*	*Pre-test mean control*	*Std.*	*n*	*Sig-two tailed*
Year 6 (Tutee) School A	20.02	5.94	46	23.51	5.4	49	.004
Year 7 (Tutee) School B	11.54	5.09	35	13.69	5.49	26	ns

Table 11. Pre/post-test gain differences

Student Performance	Quasi Pre-post Experimental Design					
	Peer Tutoring Est. Mean	Std. Error	Control Est. Mean	Std. Error	MSE	ES
Year 6 (Tutee)School A	25.90*(n=46)*	.485	22.95*(n=49)*	.470	10.32	.92**
Year 7 (Tutee)School B	15.37*(n=35)*	.702	14.46*(n=26)*	.817	16.93	.22
Year 8 (Tutee)School C	Singe group pre-post experimental design					
Pre-test mean	Sd	n	Post-test mean	Sd	n	ES
8.50	2.81	44	11.45	4.46	44	.79**

** *p<.001.*

Testing for classroom differences within groups: this was conducted to investigate whether a particular classroom showed to be performing much lower or much higher than other classroom, which would have inflated the overall findings by increasing or reducing the total group mean. When testing for class effect via an ANOVA Residual Analysis of Quadratic Regression slope for school A there was a significant main effect, F(3, 92)=5.34, p=.002. However, post-hoc comparison showed that the significant main effect was mainly evident between the treatment and the control classes, as oppose to classrooms within the treatment group, which would have inflated the findings.

Also, when testing for class effect in school B via an ANOVA of Residuals from a Cubic Regression slope, there was no main significant effect on the performance scores of different classes, F(3,57)=.87, (p=.46.)

No such analysis was undertaken for school C, since the design was weak to support any claims regarding the impact of ICAT.

4.Discussion

This study has tested a new form of cross age peer tutoring intervention, influenced mainly by Social Interdependence Theory. In order to compensate for design shortfalls, two additional safeguards were applied: regression to the mean analysis and investigations on classroom effects were carried out.

The findings for school C can be viewed with scepticism due to the lack of a treatment group, in other words the gain could be simply due to the maturation effect.

In terms of the findings for school A, the findings from this study correspond with those by Rorbeck, et al., (2003) and Leung (2014), that peer learning interventions work better on younger ages. With the younger school also having a better implementation overall. Apart from a high implementation of ICAT, at least three additional factors can be discussed to help understand the findings for school A:

Firstly, ICAT combines within it elements of the most effective peer tutoring interventions. Hence the first explanation would be the peer tutoring framework adopted here, which was informed by the 'what works' literature.

Secondly, unknown research design issues steaming from the weak design could have also biased the findings, meta-analyses have shown that in peer tutoring the chosen research design influences the magnitude of the ES (Zeneli, et al., 2016). Specifically, a) issues such as the Hawthorne effect, which is the case with short interventions, b) student and teacher demoralisation in the control groups, and c) teachers in the experimental group teaching to the test considering that all the school were keen to see peer tutoring work.

The final reason why the effect size was high in school A is that the instrument was developed to evaluate the topics covered during the ICAT intervention. If a national test was used, this would not have captured the true impact of the intervention and the effect size would have been much smaller. The high effect size in school A is consistent with peer tutoring interventions in mathematics using research made/modified instruments for the context, (Fantuzzo, Polite & Grayson, 1990; Ginsburg-Block, & Fantuzzo, 1998; Menesses & Gresham, 2009).
This research consists of major methodical limitations such as, lacking random sample selection, random participants allocation, a short time-frame, a weak instruments and many other elements necessary for a strong research which measures impact as identified by Zeneli (2016). Therefore the findings cannot be generalised to any population.

Largely, there is a need for a strong external validity design, such as *clustered*, at the school level as conducted by Tymms, et al., (2011), *stratified* random selection (Torchim, 2012); and strong internal validity, such as *blocked* randomisation allocation of schools to groups (Lachin, Matts & Wei, 1988), as well as *established instruments* (Slavin & Maden, 2008, 2011). Such designs are the way forward to shed more light on the effectiveness of the intervention developed here; as well as other education interventions in general. Also, if the intervention is lengthy enough, research needs to measure tutors' attainments, especially when considering the idea that tutors learn more than tutees (Fitz-Gibbon, 1985), which has also been established by one of the most cross-age peer tutoring robust trials (Tymms, et al., 2011). In terms of method, a better design would have been to have a four way factorial design; one group concentrating on cross-age elements, emphasising *learning processes*, one on interdependent elements, emphasising *motivational processes*, the third group emphasising both *learning* and *motivational* and a control group.

Furthermore, future research needs to go one step further and also incorporate reward interdependence. Especially when considering that the latest meta-analysis in peer tutoring has shown that studies which incorporate *reward interdependence* as chosen by students in peer tutoring provide that highest effect size (Rohrbeck, et al., 2003). This would further strengthen a cross-age peer tutoring framework and increase the level of interdependence and consequently motivation within peer tutoring pairs or triads.

Acknowledgments:
This is an ESRC funded research, DTC grant number: ES/I007296/1

References

Allan, R. (2008). *MathsLinks 1; Year 7*. Oxford, Oxford University Press.

Allen, V. L. & Feldman, R. S. (1976). Studies on the role of tutor. In Allen, V. L. (ed.) *Students as Teachers: Theory and Research on Tutoring* (pp. 113-128). Academic Press, INC: London.

Argyle, M. (1976). Social Skills Theory. In Vernon L. Allen (eds.) *Students as Teachers: Theory and Research on Tutoring*. Academic Press, INC: London.

Arreaga-Mayer, Carmen., Terry, B.J, & Greenwood, C.R. (1998). Classwide Peer Tutoring. In Topping, K.J., & Ehly (Eds.). *Peer Assisted Learning* (pp.105-121). London UK: Lawrence Erlbaum.

Barron. A-M., & Foot. H. (1991). Peer tutoring and tutor training. *Educational Research*, 33(3), 174-185.

Cohen, P.A., Kulik, J.A., & Kulik, C-L.C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*, 237–48.

DePaulo. B.M., Tang. J., Webb. W., Hoover. C., Marsh. K. and Litowitz. C., (1989). Age Differences in Reactions to Help in a Peer Tutoring Context. *Child Development*, *60*, (2), 423-439.

Fantuzzo, J.W., King, J.A., & Heller, L.R. (1992). Effects of reciprocal peer tutoring on mathematics and school adjustment: A component analysis. *Journal of Educational Psychology, 84*(3), 331-339.

Fantuzzo. J.W., Polite, K., Grayson, N. (1990). An evaluation of reciprocal peer tutoring across elementary school settings. *Journal of School Psychology, 28*(4), 309-326.

Fitz-Gibbon, C.T. (1985). Peer tutoring projects, social education improves achievement. *Journal of the National Organisation for Initiatives in Social Education*, *4*(2), 5-10.

Fitz-Gibbon, C.T. (1992). Empower and Monitor: The Em Algorithm for the creation of Effective schools. In Bashi, J., & Sass, Z. (Eds.). *School Effectiveness and Improvement* (pp 257-266.). Jerusalem: The Magnes Press.

Fitz-Gibbon, C. (2000). Cross-age Tutoring: Should it be required in order to reduce social exclusion? In Walraven, G., Parsons, C., Veen, D., & Day, C. (eds.) *Combating Social Exclusion through Education: Laissez-faire, Authoritarianism or Third Way?* (pp. 307-315). Leuven: Garant.

Fuchs, S.L., Fuchs, D., Karns, K., Hamlett, C.L., Katzaroff, M., & Dutka, S. (1998). Comparisons among individual and cooperative performance assessments and other measures of mathematics competence. *The Elementary School Journal, 99*(1), 23-51.

Ginsburg-Block, M.D., & Fantuzzo, J.W. (1998). An evaluation of the relative effectiveness of NCTM standards-based interventions for low-achieving urban elementary students. *Journal of Education Psychology, 90*(3), 560-569.

Ginsburg-Block, M., Rohrbeck, B., Cynthia, A., Fantuzzo, J. W. (2006). A meta-Analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*. 98(4), 732-749.

Harrison, G. Von & Cohen, A. M. (1971). Empirical Validation of Tutor-Training Procedures. *Annual Meeting of the California Education Association*, San Diego, California, April 1971. Retrieved November, 10, 2012, from http://www.eric.ed.gov/ERICWebPortal/search/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=ED058283&ERICExtSearch_SearchType_0=no&accno=ED058283.

Howell, K.W. (1978). Using peers in drill-type instruction. *The Journal of Experimental Education, 46*(3), 52-56.

Johnson, D. W. (1990). *Reaching Out: Interpersonal effectiveness and self-actualisation.* Prentice Hall: New Jersey.

Johnson, D.W., & Johnson, R.T. (1987). *Learning together & alone: Cooperative, competitive & individualistic learning.* New Jersey: Prentice-Hall, Inc.

Johnson, D.W., & Johnson R, T. (1989). *Cooperation and Competition: Theory and Research*. Edina, MN: Interaction Book Company.

Johnson, D.W., Johnson, R.T., Holubec, E.J., & Roy. P. (1984). Circle of Learning: Cooperation in the classroom. Alexandria, VA: Associations for supervision and Curriculum Development.

Kramarski, B., & Mevarech, Z.R. (2003). Enhancing Mathematical Reasoning in the Classroom: The Effects of Cooperative Learning and Metacognitive Training. American Educational Research Journal, 40(1), 281-310.

Lachin, J.M., Matts, J.P., & Wei, L.J. (1988). Randomization in Clinical Trials: Conclusions and Recommendations. *Control Clinical Trials, 9*, 365–74.

Leung, K, C. (2014). Preliminary Empirical Model of Crucial Determinant of Best Practice for Peer Tutoring on Academic Achievement. *Journal of Educational Psychology,* 107(2), 558-579.

Menses, K.F. & Gresham, F.M. (2009). Relative efficacy of reciprocal and nonreciprocal peer tutoring: For students at-risk for academic failure. *School Psychology Quarterly, 24* (4), 266-275.

Miller, D., Topping, K. J., & Thurston, A. (2010). Peer Tutoring in reading: The effects of role and organization on two dimensions of self-esteem. *British Journal of Educational Psychology*, 80, 417-433.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R.(2003). Peer-assisted learning interventions with elementary school students:A meta-analytic review. *Journal of Educational Psychology*, 95(2), 240-257.

Roscoe, R.D., and M.T.H. Chi. (2007). Understanding Tutor Learning: Knowledge-Building and Knowledge-Telling in Peer Tutors' Explanations and Questions, *Review of Educational Research*, 77(4) 534-574.

Roseth, C. J., Johnson, D.W., & Johnson, R.T. (2008). Promoting early adolescents, achievement and peer relationships: The effects of cooperative, competitive, and individualistic goal structures. *Psychology Bulletin*, 134 (2), 223-246.

Sharpley, A.M., & C.F. Sharpley. (1981). Peer tutoring: A review of the literature. *Collected Original Resources in Education,* 5(3), C-11.

Slavin, R. E., & Madden. N. A. (2008). Understanding bias due to measures inherent to treatments in systematic reviews in Education. Retrieved July10th, 2012, From: http://www.bestevidence.org/methods/understand_bias_Mar_2008.pdf.

Slavin, R.E., & Madden, N. A. (2011). Measures inherent to treatments in program effectiveness reviews. *Journal of Research on Educational Effectiveness, 4*(4), 370-380.

Thurston, A. (2014). Manual to help develop peer tutoring in the school classroom. Retrieved January 15th , 2015 from: http://pure.qub.ac.uk/portal/files/12775117/LT_Mathematics_PackVQUB.pdf.

Topping, K. J. (2001). *Peer assisted learning: A practical guide for teachers*. Cambridge: Brookline Books.

Topping, K. J., Campbell, J., Douglas, W. & Smith, A. J. (2003). Cross-age peer tutoring in mathematics with 7 & 11 year olds: Influence on mathematical vocabulary, strategic dialogue and self-concept. *Educational Research,* 45, 287-308.

Topping, K.J., & Ehly (1998). *Peer Assisted Learning*. London UK: Lawrence Erlbaum.

Topping, K. J., & Ehly, S. W. (2001). Peer assisted learning: A framework for consultation. *Journal of Educational and Psychological Consultation*, 12, 113– 132.

Topping, K. J., Miller, D., Murray, P., & Conlin, N. (2011). Implementation process in peer tutoring of mathematics. *Educational Psychology*, 31, 575-593.

Trochim, W. (2012), Design, the research method knowledge base. Retrieved January, 14th, 2012 from: http://www.socialresearchmethods.net/kb/design.php.

Tymms, P. (2011). Evidence? The impact of large-scale reform in England. *Zeitschrift für Erziehungswissenschaft,* 13, 105-115.

Tymms, P. B. & Merrell, C. (2015). Cross-age Peer Learning. *Better Evidence-based Education,* 7(1), 18-19.

Tymms, P., Merrell, C., Thurston, A., Andor, J., Topping, K.J., & Miller, D.J.(2011). Improving attainment across a whole district: Peer tutoring in a randomised controlled trial. *School Effectiveness and School Improvement, 22* (3), 265-289.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research,* 13, 21-40.

Webb, N. M., Troper, J., & Fall, J. R. (1992). Effective helping behaviour in cooperatives mall groups. Paper presented at the Annual Meeting of the American Educational Research Association, April, San Francisco.

Zeneli, M. (2015). Developing, Testing and Interpreting a Cross Age Peer Tutoring Intervention for Mathematics: Social Interdependence, Systematic Reviews and an Empirical Study. *Doctoral Thesis, Durham University.* www.ethesis.dur.ac.uk.

Zeneli, M. (2016). Guidance, Regulator and a Rating System for Trials/Experiments in Social Sciences: Consolidated Conduction, Reporting and Evaluation of Trials Entity (CONCRETE). *Online Journal of Education.* February, 2016, 1-20.

Zeneli, M., Tymms. P., Bolden, D. (2016). The impact of interdependent cross-age peer tutoring on social and mathematics self-concepts. *International Journal of Psychology and Education-al Studies,* 3(2), 1-13.

Zeneli, M, Thurston, A and Roseth, C., (2016). The influence of experimental design on the magnitude of the effect size -peer tutoring for elementary, middle and high school settings: A meta-analysis. International Journal of Educational Research.

Zeneli, M and Tymms, P., (2015). A Review of Peer Tutoring Interventions and SocialInterdependence Characteristics. *International Journal for Cross-Disciplinary Subjects in Education,* 5(2), 2504-2510.