The influence of social relationships on outcomes in mathematics when using peer tutoring in elementary school

Thurston, A., Roseth, C., Chiang, T-H., Burns, V., & Topping, K. (2020). The influence of social relationships on outcomes in mathematics when using peer tutoring in elementary school. *International Journal of Educational Research Open*, [100004]. https://doi.org/10.1016/j.ijedro.2020.100004

Published in:
International Journal of Educational Research Open

Document Version:
Publisher's PDF, also known as Version of record

International Journal of Educational Research Open xxx (xxxx) xxx

Contents lists available at ScienceDirect

International Journal of Educational Research Open

journal homepage: www.elsevier.com/locate/ijedro

The influence of social relationships on outcomes in mathematics when using peer tutoring in elementary school

Allen Thurston[a,b,*], Cary Roseth[c], Tien-Hui Chiang[a], Victoria Burns[d], Keith J. Topping[e]

[a] Zhengzhou University, China
[b] Queen's University Belfast, 20 College Green, Belfast BT7 1LN, UK
[c] Michigan State University, UK
[d] Durham University, UK
[e] University of Dundee, UK

ARTICLE INFO

Keywords:
Peer tutoring
Elementary mathematics
Pupil status
Additional support needs
Social inclusion

ABSTRACT

Reciprocal peer tutoring in mathematics was conducted with 487, ten to twelve year-old students from 20 elementary schools in three different school districts. The peer tutoring technique, a form of paired mathematics, placed specific emphasis on mediation through strategic metacognition between tutor and tutee. Student mathematics attainment significantly increased using this technique (Effect Size=+0.43). Student perception of the social status of their tutoring partner influenced attainment outcomes. Greatest mathematics attainment gains were predicted by having a higher opinion of the cognitive ability of students' mathematics tutoring partner and by having a mathematics tutoring partner that you believed was less popular. After peer tutoring, students showed increased social relationships in and out of school. Gains in social relationships were indicative of a more inclusive classroom being developed. The implications for theory, policy, practice and future research are discussed.

Introduction

Peer tutoring is a structured form of peer learning. It relies on constructivist approaches to learning and is based on the idea that knowledge acquisition occurs as a social activity (De Lisi & Golbeck, 1999). It is widely reported to have beneficial effects on learning (for example Ginsburgh-Block, Rohrbeck & Fantuzzo, 2006; Rohrbeck, Ginsburgh-Block, Fantuzzo & Miller, 2003; Topping, Kearney, McGee & Pugh, 2004). A meta-analytic review of peer learning reported large effect sizes around interventions which promote cognitive growth in mainstream elementary schools (Rohrbeck et al., 2003). Peer tutoring is characterized by specific role taking as either tutor or tutee by students, with a high focus on curriculum content and with clearly procedures for interaction, in which participants receive generic and/or specific training. Some peer tutoring methods scaffold the interaction with structured materials, whist others prescribe structured interactive behaviours that can be effectively applied to any materials of interest.

Literature review

Traditional peer tutoring in maths involves fixed roles, with one student being the tutor and the other being the tutee. The assumption is that the tutor is able to complete mathematics questions independently and, consequently, also capable of guiding a tutee towards the answer. A randomized controlled trial of fixed-role peer tutoring in maths in 80 schools in Scotland reported evidence of some positive effects during a two-year implementation (Tymms et al., 2011; Tymms, Merrell, Andor, Topping & Thurston, 2009). Specifically, cross-age, cross-ability fixed-role peer tutoring in 129 elementary schools had a positive effect size of +0.20 on mathematics test scores (although this effect size failed to be statistically significant), whilst *same*-age, cross-ability fixed-role tutoring had zero effect. One possible explanation for this mixed finding is that it same-age tutors are not consistently able to guide their peers towards solutions in math, even when they have different ability levels (i.e., cross-ability tutoring). A randomised controlled trial of fixed role, cross age tutoring in elementary school mathematics failed to detect effects (Lloyd et al., 2018).

An alternative to the fixed-role approach is reciprocal peer tutoring, in which each member of a dyad alternates in the role as tutor or tutee. Reciprocal peer tutoring was originally designed for pairs of low-achieving, urban, elementary school students with comparable ability levels (Fantuzzo, King & Heller, 1992) with the objective of keeping both tutor and tutee actively engaged in the academic process. Before beginning reciprocal peer learning, dyads receive training in setting joint

* Corresponding author at: Queen's University Belfast, 20 College Green, Belfast BT7 1LN, UK.
E-mail address: a.thurston@qub.ac.uk (A. Thurston).

https://doi.org/10.1016/j.ijedro.2020.100004
Received 25 May 2020; Received in revised form 29 May 2020; Accepted 3 June 2020
Available online xxx

goals as well as time limits for achieving these goals. Dyads also receive rewards if they achieve the goals that they set such as classroom certificates or other forms of recognition of achievement. Monitoring by the teacher ensures that realistic goals are set and the difficultly of the work is increased as performance is enhanced.

Research suggests that reciprocal peer tutoring results in increased attainment gains when compared to fixed-role peer tutoring (Chapman, 1998), presumably because the reciprocal approach provides students with the opportunity to act both as tutors and tutees. In addition, peer tutoring emphasizes procedural support rather content expertise by having tutors ensure that tutees follows specific steps in solving a mathematics question. In this way, tutee's steps are mediated in reciprocal peer tutoring, which allows power and learning benefits to be more evenly shared between tutor and tutee. Reciprocal peer tutoring may also involve more cooperative interaction than fixed-role tutoring because switching roles ensures that students depend on each other, rather than the tutee depending entirely on the tutor.

One other explanation for why fixed-role peer tutoring may be less effective among same-age peers is that status differences associated with being the tutor (high status) and tutee (low status) interfere with teaching and learning (Sharan, 1980). Supporting this view, a study of peer tutoring study involving 104 same-sex pairs of twelve-year-old students reported that satisfaction with learning and perceived achievement were greater when students served as the tutor than the tutee (Rosen, Powell, Schubot & Rollins, 1978). Research involving 112 nine-year-old students in a modern languages setting also suggests that methods of selecting the roles of tutor or tutee can influence outcomes. Specifically, when students were assigned the role of tutor or tutees *without* a rationale, both tutors' and tutees' attitudes towards the learning process improved. However, when intellectual characteristics were used as the selection criteria for tutors, attitudes towards learning improved among the tutors but not tutees (Bierman & Furman, 1981). A study in reading reported positive effects for tutors were about three fold when compared to gain for tutees (Thurston, Cockerill, Craig & Bersh, 2018). Taken together, these findings suggest that status congruence rather than the role of tutor or tutee per se may be the mechanism by which outcomes diverge amongst tutors and tutees.

At least three other studies suggest that social status affects the results of peer tutoring. First, in a study involving 24 children with learning difficulties, Thomson (1993) reported that tutors with lower social status were not only less effective than higher-status tutor, but also received more negative peer evaluations. Second, in a study of 53 dyads of 9- to 11-year-old children, Cole et al. paired typically developing students with students with disabilities in both play or academic tutoring contexts (Cole, Vandercook & Rynders, 1988). Results indicated that mainstream children who tutored students with disabilities tended to use less appropriate levels of play, less cooperative play and give less positive reinforcement to their tutee than mainstream children paired with students with disabilities in a free place setting. Finally, third, in a study looking at disabled preschool children in 34 classes, Odom and colleagues (Odom et al., 2007) found that children who received low sociometric scores from classmates had less positive peer interactions. Thus, all three of these studies suggest that status differences diminish the quality of peer relations.

Importantly, research suggests that some forms of peer learning influence social relationships positively. For example, a study involving 154, 5th- and 6th-grade students in an upper-middle class suburb of a large mid-western city, USA, found that co-operative learning in science promoted more positive relationships between boys and girls and between mainstream students and students with disabilities (Johnson, Johnson, Scott & Ramole, 1985). In a quasi-experimental study of 70 pupils from one elementary school in the mid-west, USA, it was also reported that students with disabilities made attainment gains equal to their mainstream classmates, and greater than control students, when they engaged in activity-based science learning with peer support (Mastropieri et al., 1998). Finally, a study involving 575, 9 to 12 year-old pupils engaged in collaborative learning in science reported that work and play relationships were significantly improved by engaging in group work in classes (Tolmie et al., 2010).

Meta-analytic evidence also indicates that students' social and academic gains are positively correlated with each other when students work together cooperatively. A meta-analysis of 148 studies from 11 countries indicated that for middle-grades students that academic achievement was strongly related to interpersonal perception (Roseth, Johnson & Johnson, 2008). In a meta-analysis of 36 peer learning studies in elementary schools Ginsburgh-Block et al. (2006) also reported that both social and self-concept outcomes were positively correlated with academic outcomes (Pearson's $r = 0.50$, $n = 20$, $p < .01$). What remains unclear from this research is whether positive peer relationship are a pre-condition for achievement, a consequence of achievement or a separate outcome than achievement. Given the aforementioned mixed results for students with disabilities, it also remains unclear how social and academic outcomes relate among students with disabilities. Thus, the present study contributes to the literature by examining both the predictive nature of social relationships and the extent to which findings vary for tutoring students with and without identified additional support needs in mainstream settings. Our specific research questions are:

- To what extent does tutor status, in terms of the tutees anticipated working relationship with their peer partner, and their perception of the cognitive ability, physical fitness, behaviours and popularity of their peer partner, predict attainment outcomes during reciprocal role peer tutoring?
- How does use of cooperative learning, in the form of reciprocal role peer tutoring, affect student perceptions of the cognitive ability, physical fitness, behaviours and popularity of their peer partner and how do these factors interact with pre and post-test social relationships regarding who children like to play with at break time?

Methodology

Design

A pre-post design was employed. This design recorded changes and tracked outcome measures and variables that could be attributable to peer tutoring as tutoring proceeded in two mathematics topics. The first tutoring topic was Number and Measurement, and the second focused on Data Handling. The intention of the research design was to relate process measures to products or outcomes. In that sense each student was serving as their own control.

Sample

The sample was composed of 485, ten, eleven and twelve year-old students from 20 elementary schools (mean age 10.76 years-old, *SD* 0.62 years). The sample comprised 256 male and 229 female students). 21.5% of students ($n = 105$) in the sample were claiming a Free School Meals (FSM) due to low socio-economic status. In the United Kingdom, context previous analysis has reported that being in receipt of a FSM describes students in the bottom quartile of income distribution (Hobbs & Vignoles, 2007). The ethnicity of the sample was 477 Caucasian Scottish, 4 Scottish Asian, 1 Caucasian French, 2 Caucasian Polish students and one black Nigerian. 77 pupils were reported to have a 'Statement of Additional Support Need'. In Scotland this meant that they had a specified additional support need in school that was documented and that a plan was in place to ensure that these addition support needs would be met in their classroom. Data from 53 of these pupils will be reported in this manuscript. Of these pupils the breakdown for how the school described their additional support need was as follows. 18 pupils non-specified reading difficulties, 7 diagnosed with dyslexia, 7 having English as an additional language, 6 with behavioural issues, 4 diagnosed with Attention Deficit Hyperactivity Disorder, 3 with non-specified numeracy issues, 3 with speech difficulties, 2 Asperger's Syndrome 1 Fragile-X, 1 hearing impaired, 1 autistic. Mean class size was 23.8 students (*SD* 3.91). The sample size was selected to allow the multi-level models to be run at classroom and individual levels. The sample was selected from three school districts. School districts were selected to be within 60 miles of the research base. Schools within these school districts were provided with information about the project and asked to express interest to participate in this research by completing a postal questionnaire.

Schools were then included or excluded in the sample dependant upon whether they fitted pre-defined selection criteria. Schools were selected on the basis that they (1) had a class available of 10–12 year old students; (2) had a class size of between 15 and 33 students; (3) had children from only one year group; (4) were capable of supporting same sex and mixed sex tutoring pairs; (5) had a percentage of free school meals in their class of not greater than 27.1% and not less than 7.1% (+/− 10% within the Scottish average for 2007 (Scottish Government, 2007); (6) were willing to take part in the project, commit to attending continuing professional development (CPD) days for relevant staff and to implement the tutoring technique in their classroom. 39 schools responded to the questionnaire. Of these 26 fitted the selection criteria. Invitations to participate were sent to all these schools. 21 schools accepted the invitation to participate in the project, but one school subsequently withdrew early in the project due to the personal circumstances of the classroom teacher.

Informed ethical consent was obtained from all participants in the study. All participants were provided with an information sheet setting out the aims of the study and how data would be stored and used. Ethical approval was obtained from the ethics committee of the Stirling Institute of Education at University of Stirling, and from appropriate personnel in the school districts in which the research took place.

Intervention

The intervention was a 16-week peer tutoring initiative in elementary school mathematics. Students spent one hour per week, for two eight-week blocks on structured mathematics peer tutoring activities (16 weeks in total). Students undertake the role of peer tutor in one topic and peer tutee in the other. Students were matched on the basis of prior attainment. The teacher was asked to rank order their class from 'highest mathematics attainment' to 'lowest mathematics attainment'. The teacher then paired the two best attaining students, the second best attaining pair of students in mathematics, the third best attaining pair and so on until each student was matched with a pair of similar mathematics attainment. Teachers were allowed to use their discretion to alter pairs if they felt that a particular pairing would result in ineffective working for social reasons. In addition 8 out of the 20 classes had an uneven number of students in the class. This meant that one triad was formed in each of these classes.

Paired Mathematics is a method of learning in mathematics in which discussion between two students (tutor and tutee) is used to solve the mathematics question. The role of the tutor is to provide support to the tutee and mediate the learning processes for the tutee. In order to do this the tutor will try to ensure that the tutee attempts the mathematics question using a structured approach and employs metacognitive strategies to approach and solve the question posed. It is the job of the tutor to keep the tutee working within this structured framework. It is the job of the tutee to do the actual working out to arrive at an answer to the mathematics question. Paired mathematics focuses on pairs of students working together and solving mathematics questions in three main steps:

- Understanding the question
- Finding an answer to the question
- Finishing the question by asking themselves what have they done and how it links to things they have done in the past

To facilitate this discussion, the pairs used the following strategies. Firstly they made sure that they understood the question. This involved reading the question, identifying what the questions was asking them to do and deciding what the best approach would be to solving the question. The pair then worked to solve the question. During this process, the tutee would think aloud and talk to the tutor who would give praise and encouragement as appropriate. The tutor would also ask the tutee questions at each step ensuring that the tutee could explain and justify the approaches they planned to take. After finishing the question the pair checked their answer and reflected on their work by asking themselves what have they done and how it linked to things they had done in the past.

The form of peer tutoring used in Paired Mathematics can trace its lineage back to a technique that was developed from a meeting of USA maths teachers. About 160 teachers were invited to brainstorm what they perceived to be effective teaching and coaching behaviours they utilised when they were working one-to-one with a student. These were summarized and sorted. The behaviours that were likely to survive transfer to interactions involving peer tutoring were developed into a procedure known as DUOLOG maths. This pedagogy was developed and used in a project in Fife by Keith Topping, David Miller (both University of Dundee) and Allen Thurston (Queen's University Belfast) in a randomized trial (Tymms et al., 2011).

Continuing professional development

Teachers from the selected schools attended three CPD days to train them in using peer-tutoring techniques. The CPD covered broad advice on how to use the technique, the topics of mathematics that the technique was to be utilised during and provided advanced development in creating effective peer tutoring partnerships in the classroom. This included specific advice on how to enhance social interactions and communication skills within pairs of students.

Instrumentation

Mathematics attainment: Criterion referenced attainment tests in mathematics were developed from tests used in The Scottish Assessment & Achievement Programme Survey 2004. This was a nationally available instrument, independently designed to be aligned to the Scottish Curriculum by the Scottish Qualifications Authority. Good reliability and validity was reported for this instrument. Cronbach alpha values were reported to be between 0.7 and 0.9 when the attainment test was used by the Scottish Government with a sample of 2345, ten to twelve year-old Scottish students from the similar school class year-groups used in the current project (Scottish Government, 2004). The instrument was scored out of 50. Marking schedules were agreed before any tests were marked and clear protocols were established in respect of the administration and marking of the attainment tests. Marking schedules were developed using the framework for assessment and marking utilised by the Scottish Assessment & Achievement Programme Survey 2004 (Scottish Government, 2004).

Attitudes to my maths partner instrument: The 'Attitudes to my maths partner' instrument was designed to ascertain the degree to which attitudes to a social status of a maths partner may influence outcomes of peer tutoring in maths. The instrument was designed to measure:

- How tutors expected to work with their maths partner.
- The perception of their maths partners cognitive ability in mathematics.
- The perception of level of physical fitness and status of their peer partner.
- The perception of the behaviour standards of their peer partner.
- Their perception of the popularity of their peer partner.

In order to do this a 20 item questionnaire was designed with four questions being asked on each aspect of tutor status. Questions were pre-

Table 1
Factor loadings, eigenvalues and factor reliability from Attitudes to maths scale taken from pattern matrix

Pattern Matrix	Rotated Factor loadings				
Item	Expected working	Ability	Physical Fitness	Behaviour	Popularity
My maths partner and I will work well together	**.83**	−.09	−.06	.15	.05
I will like working with my maths partner	**.83**	−.03	−.05	.07	−.08
I look forwards to working with my maths partner	**.82**	−.04	−.13	.11	−.00
People like my maths partner	**.40**	−.19	.06	−.23	−.39
My maths partner is good at maths	.18	**-.84**	.11	.00	.06
My maths partner is brainy	.11	**-.82**	.03	−.05	−.03
My maths partner is clever	.12	**-.81**	−.01	−.02	.02
Mathematics is difficult for my maths partner	.30	**.76**	.14	−.08	.02
I don't think my maths partner is very fit	−.06	.040	**.83**	−.13	−.10
My maths partner is not good at PE	.07	−.03	**.82**	−.11	.01
My maths partner can run fast	.07	−.04	**-.60**	−.25	−.21
My maths partner is good at sports	.19	−.03	**-.54**	−.32	−.13
I do not think m y maths partner behaves very well	−.12	−.02	.07	**-.74**	.23
My teacher thinks my maths partner is well behaved	.27	−.16	−.10	**.63**	.06
My maths partner has strong muscles	.22	−.06	−.30	**-.45**	−.12
My maths partner isn't good at school work	−.02	.32	.02	−.37	.30
My maths partner tends to play by themselves in the playground	.14	−.01	.04	−.11	**.84**
My maths partner does not have many friends	−.10	−.03	.04	−.06	**.76**
People in my class are friends with my maths partner	.39	−.03	−.10	−.07	**-.40**
My maths partner is popular	.23	−.22	−.08	−.30	−.39
Eigenvalues	7.32	2.11	1.36	1.23	1.01
% of variance	36.59	10.53	6.80	6.13	5.03
α	.87	.83	.77	.47	.74

sented in the form of a statement about their peer partner. Respondents were asked to indicate whether they agreed (5 points) to disagreed (1 point) with the statement about their peer partner on a five point Likert scale. Only the poles were marked on the Likert scale and 13 questions were worded such that 'agree' would be a positive response regarding their peer partner and 7 were worded with reverse polarity such that 'disagree' would indicate a positive response about their peer partner. This instrument was newly designed specifically for this research. The principal component analysis is presented below.

In addition to the attitudinal scales, the students were also asked to report who in their class, they liked to play with at break time. This was done by providing each child with a class list, and asking them to mark the children that they liked to play with at break. This was completed at both pre and post-test.

A principal component analysis (PCA) was conducted on the 20 items of the Attitude to my Maths Partner time-one scale with oblique rotation due to the expectation than the factors would be related. The Kaiser–Meyer–Olkin measure verified the sampling adequacy for the analysis, KMO=0.90 and all KMO values for individual items were >0.70, which is well above the acceptable limit of 0.50 (Field, 2009). Bartlett's test of sphericity, χ^2 (190) =4025.98, p <0.001, indicated that correlations between items were sufficiently large for PCA. An initial analysis was run to obtain eigenvalues for each component in the data. Five components had eigenvalues over Kaiser's criterion of 1 and in combination explained 65.08% of the variance. (The scree plot was ambiguous and showed inflections that would justify retaining 2, 4 or 6 components.) Given the sample size ($N = 418$) and that Kaiser's criterion is accurate with sample sizes over 250 if the average communality is greater than 0.6 (0.65 in this case), it was decided to retain 5 components. Table 1 below shows the factor loadings after rotation from the pattern matrix and Table 2 shows the factor loadings after rotation from the structure matrix. The items taken as belong to each factor are highlighted in yellow. Bold fonts shows loadings of >0.4.

Interpreting from the pattern matrix (Table 1), factor 1 (Expected working) included four items which focused mostly on anticipated working with partner. Factor 2 (Ability) included four items all to do with partner's academic ability. Factor 3 (Physical fitness) included four items which were to do with physical ability e.g. sports, running. Factor 4 (Behaviour) included 3 items which loaded reasonably strongly onto the factor and one with a loading of below 0.4. The two strongest loading items in this factor were to do with view of partner's behaviour. The third item 'my maths partner has strong muscles' did not load well. The weak loading item 'My maths's partner isn't good at school work' was difficult to interpret in this factor. Factor 5 (Popularity) included 4 items all concerned with sociability and popularity of maths partner. One item 'My maths partner is popular' loaded weakly (0.39) onto this factor. However, reliability was not improved by removing any items in this factor.

A structure matrix (Table 2) was used to draw out the factors, placing each item in the factor it loaded most highly on to. It was decided to utilise this structure matrix to construct the final sub-scales. Items loading into each factor were summed to create numeric sub-totals for each sub-scale. Data indicated that the same general factor descriptions still occurred, but with less rogue items. Factor 1 (Expected working) remained the same. Factor 2 (ability) now contained 5 items with 'My math's partner isn't good at school work' in addition to the previous items. This item seemed more logically placed here despite the fact that ironically there was a slight reduction in reliability. Factor 3 (physical ability) now contained 5 items with the additional item 'My math's partner has strong muscles' which again seemed more logically placed in this factor and improved the reliability slightly. Factor 4 (behaviour) contained only the 2 previous items pertaining to classroom behaviour and reliability increased to α=0.66. Factor 5 (popularity) remained unchanged. A combined Cronbach's alpha reliability statistic was calculated for the Attitudes to Maths partner scale including all 20 items α=0.90. This could not be significantly improved by deleting any items.

Results

Multi-level models were used to assess changes over time in mathematics attainment and students' attitudes toward their peer partner. In addition, multi-level models were used to examine the extent to which social factors such as tutor status and students' attitudes toward their partner predicted mathematics attainment.

Table 2
Factor loadings, eigenvalues and factor reliability from Attitudes to maths scale taken from structure matrix.

Structure Matrix	Rotated Factor loadings				
Item	Expected working	Ability	Physical Fitness	Behaviour	Popularity
I will like working with my maths partner	**.89**	**-.40**	**-.40**	.03	**-.44**
I look forwards to working with my maths partner	**.89**	**-.42**	**-.45**	.08	**-.40**
My maths partner and I will work well together	**.87**	**-.44**	−.38	.13	−.35
People like my maths partner	**.62**	**-.43**	−.32	−.23	**-.60**
My maths partner is good at maths	**.44**	**-.86**	−.19	.12	−.27
My maths partner is brainy	**.42**	**-.85**	−.27	.07	−.35
My maths partner is clever	**.43**	**-.85**	−.29	.10	−.33
Mathematics is difficult for my maths partner	−.05	**.71**	.27	−.20	.23
My maths partner isn't good at school work	−.25	**.50**	.24	**-.41**	**.42**
I don't think my maths partner is very fit	−.35	.30	**.83**	−.13	.29
My maths partner is not good at PE	−.22	.21	**.79**	−.10	.32
My maths partner can run fast	**.41**	−.28	**-.73**	−.26	**-.51**
My maths partner is good at sports	**.46**	−.26	**-.67**	−.33	**-.44**
My maths partner has strong muscles	**.43**	−.21	**-.46**	**-.45**	−.37
I do not think m y maths partner behaves very well	−.20	.24	.20	**-.72**	.28
My teacher thinks my maths partner is well behaved	.32	−.36	−.22	**.65**	−.14
My maths partner does not have many friends	**-.41**	.31	.38	−.04	**.80**
My maths partner tends to play by themselves in the playground	−.20	.27	.34	−.10	**.80**
People in my class are friends with my maths partner	**.60**	−.34	**-.42**	−.09	**-.61**
My maths partner is popular	**.52**	**-.43**	−.40	−.28	**-.60**
Eigenvalues	7.32	2.11	1.36	1.23	1.01
% of variance	36.59	10.53	6.80	6.13	5.03
α	.87	.81	.78	.66	.74

Table 3
Multi-level Modelling Analysis of Pre- Post-change in Maths and Peer Attitudes.

Model	Maths Attainment	Liking	Attitudes towards Peer Partner: Work with	Ability	Fitness	Behavior	Popularity
Fixed Effects							
Intercept	−2.36	23.26	17.61***	14.31***	19.07***	5.61*	12.78***
Age	0.20**	0.06	−0.01	0.03	−0.01	0.01	0.01
Gender	−0.25	0.83	0.15	0.35	0.01	−0.19	0.34
FSM	−2.83***	−2.67	0.19	−0.21	0.11	0.01	−0.01
ASN	−7.71***	−2.45	0.19	−1.49**	0.68	−0.25	−0.44
Class size	0.03	−0.36	0.03	−0.01	0.05	0.01	0.02
Time	2.91***	4.82*	−1.26***	−0.28	−0.94**	−0.19	−0.01
Random Effects							
School-level							
UN(1,1)	18.51**	25.52*	0.40	0.67	0	0.23	0.35
UN(2,1)	−3.74	0.85	0.40	−0.28	−0.20	−0.15	−0.19
UN(2,2)	1.77*	54.29**	0.98	0.67	0.65	0.13	0.21
Student-level							
UN(1,1)	38.90***	89.51***	11.44***	10.62***	15.27***	2.97***	6.93***
Residual	20.18***	106.45***	7.92***	8.58***	7.50***	2.58***	5.39***

Mathematics attainment

The results of multi-level modelling reported in Table 3 indicate that mathematics attainment increased over time ($B = 2.91$, $p < .001$), even as the initial levels of attainment were lower among students with FSM ($B = -2.83$, $p < .001$) and ASN ($B = -7.71$, $p < .001$). The only other predictor of initial mathematics attainment was age ($B = 0.20$, $p < .01$), as neither gender nor class size were significant. There was no evidence that the magnitude of change over time as a function of age, gender, FSM, ASN or class size (interaction results not presented).

Attitudes towards partner

The ‘attitudes to my peer partner’ instrument was administered at pre-test and at post-test. Significant and reasonably strong correlations were observed in all five sub-scales between pre and post-test attitudes:

- How tutors expected to work with their maths partner ($r = 0.59$, $n = 369$, $p < .001$).
- The perception of their maths partners cognitive ability in mathematics ($r = 0.51$, $n = 369$, $p < .001$).
- The perception of level of physical fitness and status of their peer partner ($r = 0.66$, $n = 369$, $p < .001$).
- The perception of the behaviour standards of their peer partner ($r = 0.53$, $n = 369$, $p < .001$).
- Their perception of the popularity of their peer partner ($r = 0.55$, $n = 369$, $p < .001$).

The results of multi-level modelling reported in Table 3 indicate that only two of the five attitudes toward partners changed significantly over time. Specifically, students reported that both their expectations for working well with their maths partner ($B = -1.26$, $p < .001$) and their perception of their partner's physical fitness and status ($B = -0.94$, $p < .01$) decreased significantly over time. Moreover, there was little evidence that attitudes varied as a function of other characteristics, save for the fact that students with ASN reported significantly lower perceptions of partner's cognitive ability in mathematics.

Table 4
Multi-level Modelling Analysis of Pre- Post-change in Maths Predicted by Peer Attitudes and Peer Liking

Model	1	2	3
Fixed Effects			
Intercept	−2.36	−8.17	−8.74
Age	0.20**	0.18**	0.19**
Gender	−0.25	−0.03	0.04
FSM	−2.83***	−2.45**	−2.45**
ASN	−7.71***	−7.10***	−7.02***
Class size	0.03	0.08	0.04
Time	2.91***	8.93***	10.38***
Works with		−0.01	0.01
Ability		0.40***	0.40***
Fitness		0.04	0.04
Behavior		0.07	0.11
Popularity		−0.20*	−0.20*
AttTut_Cog*Time		−0.25**	−0.23**
Liking			0.01
Liking*Time			−0.05*
Random Effects			
School-level			
UN(1,1)	18.51**	19.47**	19.25**
UN(2,1)	−3.74	−5.80	−5.72
UN(2,2)	1.77**	8.73**	8.40**
Student-level			
UN(1,1)	38.90***	35.50***	36.68***
Residual	20.18***	18.54***	18.33***

As reported in Table 4, results also indicated that two student attitudes predicted mathematics attainment. First, students' perception of their partner's ability in mathematics was associated with increased maths attainment ($B = 0.40$, $p < .001$), although a significant interaction indicated that this relation diminished at post-test ($B = -0.25$, $p < .01$). Second, students' perception of the popularity of their peer partner was associated with decreased maths attainment ($B = 0.20$, $p = .03$). There were no changes in the relation between other variables (e.g., Age, FSM, ASN) and maths attainment after accounting for students' attitudes towards their partners.

Peer liking

As reported in Table 3, peer liking (i.e., "like to play with at breaktime") increased significantly over time ($B = 4.82$, $p = .01$), and there was no evidence that initial levels of liking nor the magnitude of change varied as a function of age, gender, FSM, ASN or class size.

As reported in Table 4, results also indicated that peer liking predicted maths attainment, even after controlling for students' attitudes and demographic and classroom characteristics. However, a significant interaction between time and peer liking indicated that this relation was small in magnitude and diminished at post-test ($B = -0.05$, $p < .05$).

Discussion

Data indicated that social relationships play a complex role in academic outcomes during the peer tutoring process. On the one hand perception of the cognitive ability and popularity of a tutor partner can have a predictive affect on academic outcomes. On the other it appears that relationships out of the classroom and out of the school work as a separate and distinct outcome of peer tutoring to attainment. In a study involving 575, 9–12 year-old students in Scotland engaged in collaborative groupwork forms of peer learning it was concluded that social and cognitive gains are separate outcomes of collaborative learning (Howe et al., 2007). However, it was hypothesised that social and cognitive gains were interlinked, even if distinguishable. Social Interdependence theory links co-operative interactions to both social and achievement benefits. Roseth et al. (2008; p226) reported that *'The successful accomplishment of group members' goals results in a positive cathexis that is generalized to each other, resulting in more positive socialrelationships. A benign spiral results in which positive social relationships increase promotive interaction, which increases achievement, which increases positive cathexis, which increases positive social relationships even more, and so forth'.* Their meta-analysis of 17,000 students involved in 148 studies yielded positive effect sizes in both student attainment (*ES*=0.46) and social relationships (*ES*=0.48) for co-operative over individualistic forms of learning. They reported significant relationships between peer attraction and attainment outcomes. Findings in this study appear to being line with these findings with greater perception of the cognitive ability of the peer partner popularity significantly predicting cognitive gains. In addition, pre-test cognitive test scores in mathematics were related to pre-test sociometric connectedness. This seemed to indicate that social and attainment outcomes in classes are linked, but causality cannot be established from the present study. Such correlational relationships between cognitive outcomes and social connectedness have been reported previously (Thurston et al., 2010). There may now be reasons to design experiments to determine causality in this relationship.

Data indicated that more popular peers made worse peer partners and that there was a need to have a peer partner who was more highly thought of in terms of their cognitive ability. The popularity subscale included questions that measured the peers' perceptions of their maths partner in respect of: How many friends they had; the degree of their social interaction when playing in the playground; The number of people in the class that they were friends with, and; their general popularity. Each of these measures is a reflection of how well their maths partner mixed socially with other class mates in and out of the classroom. However, it would appear that there is a distinction between effective peer interactions on a social basis and those that are designed to promote cognitive development. The fact that a high a cognitive perception (and this was shown to be significantly related to actual cognitive ability) of the maths partner was related to more effective learning could be explained in a number of ways. Pairs were matched with students that the teacher perceived to be close in terms of cognitive status. In such a relationship it may be easier for the status that one confers on one's peer, to be conferred on oneself! If you perceive your peer to be good at math and you are at the same level, ergo you must be good at math. This may lead to a boost in self concept and confidence. Alternatively it may be that too large a gap in cognitive ability may lead to less effective tutoring. The gap between the cognitive structures of the tutee and tutor may be too great. This could lead the tutee to reject the alternative model suggested by the tutor or alternatively when acting as tutor the pair member of lower cognitive ability may not be able to mediate the learning of the tutee because they do not understand what the tutee is doing (Thurston & Topping, 2007). It would appear that if enhanced academic outcomes are the aim of teachers using mathematics peer tutoring then best advice may be to avoid grouping pairs on the basis of popularity and to form partnerships of close math ability where there is a healthy respect for the cognitive ability between the pairs.

Social relationships were enhanced for pupils during the project. Therefore, in this project peer tutoring seemed to have a dual effect enhancing both academic and social outcomes for participants. The effect sizes were not large for changes in social interaction. However, the intervention did not give rise to wholesale change in the nature of the classroom teaching and learning. It was based on one-thirty minute session per week for 16 weeks. Notwithstanding this the nature and pattern of effects was clear. The intervention was not able to provide evidence of how these social gains impacted upon academic attainment. However, working with a mathematics partner whom you perceive to be more popular was a significant predictor of academic attainment gains. Therefore, further work is required to see if the social gains would lead to additional academic gains at a later stage. In a meta-analysis of 148 cooperative learning intervention Roseth, Johnson & Johnson (2008) reported that if teacher structure classroom goals co-operatively then students would achieve better results, students would have better relationships and that higher levels of achievement would be associated with better relationships. However, their meta-analysis was not able to determine causation. This study has determined that there may be a causality of some aspects of peer relationships on peer tutoring outcomes. However, the process by which this causality may act is still not clear.

For policy makers and practitioners it would appear to be sensible to ensure that pedagogies being utilized by teachers in mathematics include careful selection of partners for peer learning. Pair pupils who have a similar cognitive ability and a good cognitive perception of each other and not to pair pupils on the basis of popularity. One of the benefits of the approach described in this manuscript is that it appeared to provide benefit to group irrespective of sex, socio-economic status or additional support needs. However, it has the added benefit of promoting social relationships for pupils who have an additional support need whilst still providing academic enhancement for all pupils. In this respect, teachers could implement the technique safe in the knowledge that no one group from those described would be unfairly disadvantaged academically, but there would be social benefits for pupils. Other forms of peer learning have reported similar effects. A study in 24 schools in Scotland reported gains in cognitive science tests and social relationships when using collaborative learning (Tolmie et al., 2010). A study in the USA mid-west reported academic and social gains when using co-operative learning in science (Johnson et al., 1985). This research identifies similar findings in the third distinct form of peer learning, peer tutoring, for the first time.

Another important factor that may have contributed to the positive effect sizes in the present study was the fact that reciprocal peer tutoring was used. An externally designed national test was utilised to measure attainment in this study and yielded positive results. Previous studies using fixed role peer tutoring in mathematics similar in nature to this study found effect sizes to be more modest in nature, and greater for tutors than tutees. Greater effect sizes were reported for cross-age, cross ability peer tutoring (Tymms et al., 2009). Tutors are reportedly perceived as having higher status than tutees within classroom settings (Sharan, 1980). Using intellectual characteristics as selection criteria (as was the case in the present study) was previously reported to enhance outcomes only for tutors (Bierman & Furman, 1981). This may account for the relatively greater benefit of reciprocal paired mathematics when compared to fixed role tutoring in the previous study.

Conclusion

The present study starts to identify the sort of pairings for peer tutoring that may be most effective to enhanced cognitive test performance. Recent studies have found cross-age peer tutoring to be less effective when used at scale (Lloyd et al., 2018), and same age cross-ability tutoring to have negligible effect (Tymms et al., 2011). It would also appear that peer tutoring, as with other forms of peer/cooperative learning can have social as well as cognitive gains (Tolmie et al., 2010). If social standing, in particular perception of cognitive ability, can influence outcomes then it may stand to reason that increases in social standing and math attainment may eventually feed forwards to promote additional cognitive gains. However, the present study was limited in terms of its length and scope. Additional work may be required to firmly establish that these factors may inter-relate with the tutoring process to influence academic outcomes.

The present study was undertaken in compulsory state education classrooms and has relied on teachers adopting a new pedagogy into their professional practice. The work has some limitations due to the very nature of 'design experiments'. These include not being able to control all variables in a classroom. Nonetheless the work reported is drawn from a wide variety of educational contexts and some of the limitations of working with authentic schools actually are a strength in providing good ecological validity of how the technique of paired mathematics with strategic metacognition manifests in a school setting. In addition, the interpretation of social relationships between students with, and without, additional support needs should be treated with caution. The analysis did not take account of clustering effects (due to the low numbers in each class/cluster) and there is a danger that other issues surrounding the sample composition of the students with additional support needs contributed to outcomes.

Further work is now required to test whether the findings in this study would survive scale-up to a larger sample. This would also include fine tuning the advice given to teachers and students to ensure that greater emphasis is given to establishing pairs on the basis of prior attainment, perception of cognitive ability and popularity. In addition, such a study could test the effects on student mathematics performance of purposeful professional development for a randomly selected group of teachers as compared to a randomly selected control group. It should be noted schools were selected after volunteering in this project and so the sampling issues may have contributed to the positive effect sizes observed in respect of mathematics attainment and enhanced implementation fidelity in participating classrooms. Sampling effects may therefore be evident in the data presented. However, as the data presented effectively self-controls on the basis of differential outcomes then this may not be as large a problem as first thought for the data presented from this study. Finally, there may be a need to undertake work to ascertain whether fixed or reciprocal role peer tutoring in mathematics is most effective.

Declaration of Competing Interest

None.

Acknowledgement

Our great thanks are due to Professor Christine Howe (University of Cambridge) who supported the development of this proposal through the Economic & Social Research Council (ESRC) Teaching & Learning Research Programme Research Mentoring Fellowship Programme, Angela O'Donnell (Rutgers University, USA) and Andrew Tolmie (Institute of Education, University of London, UK) who both acted as consultants to the project. The research project was supported by a research grant awarded by the ESRC.

References

Bierman, K. L., & Furman, W. (1981). Effects of role and assignment rationale on attitudes formed during peer tutoring. *Journal of Educational Psychology, 73*, 33–40.

Chapman, E. S. (1998). Key considerations in the design and implementation of effective peer-assisted learning programmes. In K. J. Topping, & S. Ehly (Eds.), *Peer-assisted learning* (pp. 67–84). Mahwah, NY: Erlbaum.

Cole, D. A., Vandercook, T., & Rynders, J. (1988). Comparison of two peer interaction programmes: Children with and without severe disabilities. *American Educational Research Journal, 25*(3), 415–439.

De Lisi, R., & Golbeck, S. L. (1999). Implication of Piaget's theory for peer-learning. In A. M. O'Donnell, & A. King (Eds.), *Cognitive perspectives on peer-learning*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

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

Ginsburgh-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. (2006). A meta-analytic review of social, self-concept, and behavioural outcomes of peer-assisted learning. *Journal of Educational Psychology, 98*, 732–749.

Hobbs, G., & Vignoles, A. (2007). *Is free school meal a valid proxy for socio-economic status (in schools research)?*. London: Centre for Economics of Education, London School of Economics Retrieved from http://cee.lse.ac.uk/cee%20dps/ceedp84.pdf .

Howe, C. J., Tolmie, A. K., Thurston, A., Topping, K. J., Christie, D., Livingston, K., Jessiman, E., & Donaldson, C. (2007). Group work in elementary science: organisational principles for classroom teaching. *Learning & Instruction, 17*(5), 549–563.

Johnson, R. T., Johnson, D. W., Scott, L. E., & Ramole, B. A. (1985). Effects of single–sex and mixed-sex cooperative learning on science achievement and attitudes and cross-handicap and cross-sex relationships. *Journal of Research in Science Teaching, 22*, 207–220.

Lloyd, C., Edovald, T., Morris, S., Kiss, Z., Skipp, A., & Haywood, S. (2018). Durham shared maths project. Retrieved from https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Shared_Maths.pdf (accessed 25 September 2018)

Mastropieri, M. A., Scruggs, T. E., Mantzicopoulos, P., Sturgeon, A., Goodwin, L., & Chung, S. (1998). "A place where living things affect and depend on each other": Qualitative and quantitative outcomes associated with inclusive science teaching. *Science Education, 82*, 163–179.

Odom, S. L., Zercher, C., Li, S., Marquart, J. M., Sandall, S., & Brown, W. H. (2007). Social acceptance and rejection of preschool children with disabilities: A mixed method analysis. *Journal of Educational Psychology, 98*(4), 807–823.

Rohrbeck, C. A., Ginsburgh-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*, 240–257.

Rosen, S., Powell, E. R., Schubot, D. B., & Rollins, P. (1978). Competence and tutorial role as status variables affecting peer-tutoring outcomes in public school settings. *Journal of Educational Psychology, 70*, 602–612.

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

Scottish Government (2004). *Seventh survey of mathematics 2004*. Retrieved from http://www.scotland.gov.uk/Publications/2005/10/2192247/22532 8 June 2010)

Scottish Government (2007). *School meals in Scotland 2007*. Retrieved from http://openscotland.gov.uk/Publications/2007/06/04134749/2

Sharan, S. (1980). Cooperative learning in small groups: Recent methods and effects on achievement, attitudes, and ethnic relations. *Review of Educational Research, 50*, 241–271.

Thomson, A. (1993). Communicative competence in 5- to 8-year olds with mild or moderate learning difficulties and their classroom peers: Referential and negotiation skills. *Social Development, 2*(3), 260–278.

Thurston, A., Cockerill, M., Craig, N., & Bersh, L. (2018). Assessing the differential effects of peer tutoring for tutors and tutees. In *Proceedings of the American educational research association annual gathering*.

Thurston, A., & Topping, K. J. (2007). Peer tutoring in schools: Cognitive models and organisational typography. *Journal of Cognitive Educational Psychology, 6*(3), 356–372.

Thurston, A., Topping, K. J., Tolmie, A. K., 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.

Tolmie, A. K., Topping, K. J., Christie, D., Donaldson, C., Howe, C., Jessiman, E., et al. (2010). Social effects of collaborative learning in primary schools. *Learning & Instruction, 20*, 177–191.

Topping, K. J., Kearney, M., McGee, E., & Pugh, J. (2004). Tutoring in mathematics: A generic method. *Mentoring and Tutoring, 12*, 351–368.

Tymms, P., Merrell, C., Andor, J., Topping, K. J., & Thurston, A. (2009). Improving attainment across a whole district: Peer tutoring in a randomised controlled trial. In *Proceedings of the European association of research on learning & instruction Amsterdam 24th–30th August 2009*.

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.