Education 3-13
DOI:10.1080/03004279.2019.1701513

Collaborative Group Work in Mathematics in the UK and Japan: Use of Group Thinking Measure Tests

Taro Fujita[1], Jonathan Doney[1], Ruth Flanagan[1], and Rupert Wegerif[2]

1. Graduate School of Education, University of Exeter, UK
2. Faculty of Education, Cambridge University, UK

Corresponding author: Taro Fujita; Graduate School of Education, University of Exeter, UK; 79 Heavitree Rd, Exeter EX1 2LU; t.fujita@exeter.ac.uk

Abstract: This paper reports on a study of collaborative group work in mathematics. Although collaborative group work is known as one of the important approaches in education, it is still uncertain how group thinking can be measured in various learning contexts. We used the Group Thinking Measure (GTM) test developed by Wegerif et al. (2017) alongside mathematics tests to measure group thinking and group mathematical thinking. Our participants from Japan (134 pupils, 10-12 year old) and the UK (30 pupils, 11-12 year old) schools undertook the GTM individually and then in a group of three (triad), following which, the same group also solved sets of mathematics problems. From the quantitative results we found that examining whether a group is a Value Added Group or not in their GTM scores is a useful way to identify more mathematically effective groups. From a qualitative analysis of video data of pupils' group work, we also found that successful problem solving might be due to the use of certain strategies. In conclusion, we consider that GTM can be used to indicate which groups are effective in subject areas such as mathematics.

Keyword: Collaborative group work, Group thinking measure, mathematics, UK and Japan

DOI:10.1080/03004279.2019.1701513

1. Introduction

Whilst the benefits of collaboration within the process of learning appear in the philosophical writings of Rousseau (1762) and were expanded on by Dewey (1916), educational research concerning group work did not begin in earnest until the 1970s. This was due, in part, to the rise of the influence of social constructivism (Vygotsky, 1978, Bruner, 1985) as well as technological improvements in data collection and analysis. Vygotsky's (1978) seminal work emphasised the importance of interaction with others to enhance, and provide the necessary tools for, pupils' learning. Social constructivism impacted classroom practice including layout, with furniture moved from rows to placing pupils routinely in groups around tables. Additionally, technological advancements such as audio and videotaping of group work, alongside improved systems to facilitate analysis, aided data collection and effective analysis of the impact of group work. This contributed to the growth of educational research concerning collaborative group work (Johnson and Johnson, 2000; Gillies and Ashman, 2003).

However, collaborative group work does not automatically promise effective learning, and existing studies suggest various issues to be investigated further. For example issues include: how to measure group thinking (e.g. Wegerif, et al, 2017); pupils' ways of interacting including how they engage in exploratory talk (e.g. Rabel and Wooldbrige, 2012), how they manage themselves and others (Kershner et al, 2014); how to scaffold group work (Patterson, 2016; Kazak et al, 2015a); identifying mechanisms of collective conceptual growth (e.g. Kazak, et al, 2015b); including pupils with SEN (Baines et al, 2015), difference between male/female groups (e.g. Dahl, 2018) and so on. Among many issues for effective collaborative group work, relatively little research has been done to measure group thinking, with the exception of Wegerif, et al. (2017) who have developed a test to measure group thinking based on nonverbal reasoning test questions. Their Group Thinking Measure (GTM) test involves two separate equivalent tests, one for groups and one for individuals. It was suggested that solving the test items requires pattern recognition and logical inferences, and through this they illustrated how the GTM test was used to measure individual and group thinking skills by examining how students interacted with each other when they were undertaking the test.

One of the questions derived from this study is to consider how the skills measured by the GTM test might be closely related to other types of thinking, such as mathematical thinking, and investigating to what extent collaborative group working skills in the context of mathematics is related to more general group thinking. In order to help children engage with challenging mathematical problems, collaborative group work can be an effective approach, as has been discussed (e.g. Francisco, 2013; Martin and Towers, 2015; Kazak et al, 2015a, 2015b; Dahl et al, 2018). It is therefore important to examine relationships between general group thinking and mathematical thinking in collaborative learning situations.

In this paper, we report findings from three data sets derived from two related exploratory studies in which the GTM test and mathematics test were given to 11-13 year old pupils in the UK and Japan. In the UK, the current National Curriculum states that pupils "must be assisted in making their thinking clear to themselves as well as others, and teachers should ensure that pupils build secure foundations by using discussion to probe and remedy their misconceptions" (DfE, 2014). It is necessary to develop a good model for collaborative group work in primary and lower secondary schools. We choose Japan because collaborative group work is known as one of the features in Japanese mathematics lessons (e.g. Hino, 2015).

The aims of this paper are to answer the following research questions, 'i) What are the relationships between general collaborative group thinking measured by the GTM and mathematical thinking measured by tests, and ii): What can we learn from collaborative problem solving process in GTM and maths activities?' Our intention is not to decide which country is better, but to explore the differences and similarities in order to enrich our understanding of features of successful group thinking in problem solving in mathematics in the UK and Japan, and to provide guidance for

DOI:10.1080/03004279.2019.1701513

successful collaborative group work in general, and the specific subject area of problem solving in mathematics.

2. Effective Group work

2.1 Existing research in Collaborative group work

Our concept of collaborative group work is guided by Panitz's definition of collaborative learning, respecting and highlighting people's abilities, sharing responsibilities and cooperation rather than competitions (Panitz, 1999). Group work has become the 'norm'; one national survey on group work, undertaken in the USA (Puma et al, 1993) recorded that 79% of elementary teachers and 62% of middle school teachers employed some form of cooperative learning. An increase in use was found in a study (Antil et al, 1990) where 93% of teachers, who were questioned, reported using cooperative learning with 81% of this being on a daily basis (cited by Slavin, 1995). Similar findings were recorded in the UK in the 1990s, where group work was a frequent occurrence in schools. 'The practice of organising the class into groups is common in all schools and inevitable in smaller ones' (Alexander et al, 1992, p. 29).

The educational benefits of such group work have been lauded by researchers as improving:

- Learning and conceptual development
- School achievement
- Engagement in learning
- Oracy development
- Critical and analytical thinking skills
- Motivation and attitudes
- Behaviour in class and relations with peers (Baines et al, 2009, p. 8)

There is a disconnect between the theory behind collaborative group work and the practical outworking. This may be due to the fact that the placing of pupils into

DOI:10.1080/03004279.2019.1701513

groups does not mean that effective collaborative group work is necessarily occurring: 'The fact that they are seated in groups does not necessarily mean that they are working as a group' (Alexander et al, 1992, p. 29). Recent educational research has examined the role of the pupils, in terms of the impact of their dialogic interaction on their learning. Research has been conducted into the efficacy of developing dialogic talk in group work in Science (e.g. Mercer et al, 2004; Pifarré, 2019), ICT (e.g. Wegerif and Dawes, 2001, 2004) and Literacy (e.g. Dawes, 2001, 2004; Rojas-Drummond et al, 2006; Newman, 2017).

For effective collaborative group work in problem solving ensuring that the group assist each other in developing their conceptual understanding rather than merely relying on the highest achieving pupil's answer is crucial (Pifarré and Li 2018). Language is perhaps one of the key elements of productive group learning processes. Newman (2017) notes the educational potential of collaborative talk between peers, which involves sharing perspectives, negotiating, and resolving difference (Barnes and Todd, 1977, 1997; Mercer and Littleton, 2007, Kershner et al, 2014). Mercer and Sams (2006) or Monaghan (2005) particularly consider that the role of exploratory talk (described as being critical friends to each other) and the use of explicit reasoning during problem solving, is crucial for developing understanding, in comparison with other types of talk such as disputational (being competitive or disagreeing with each other in egoistic ways) or cumulative talk (agreeing each other without constructive criticisms). Vygotskian approaches take a view that useful 'tools' such as exploratory talk should be acquired by learners, and stress the importance of setting norms or 'ground rules for talk', requiring that the views of 'all participants are sought and considered, that proposals are explicitly stated and evaluated, and that explicit agreement precedes decisions and actions, with ultimate agreement being sought' (Mercer and Howe, 2012, p. 16) in order to support the acquisition of the tools. Dahl et al. (2018) also found that the females' group exchanged their talk more effectively than the males' group, which resulted in a more productive outcome in a mathematical task.

DOI:10.1080/03004279.2019.1701513

Wegerif et al. (2017) proposed that group thinking can be measured by nonverbal reasoning tests, and thus developed the GTM test. Through their initial pilot study in implementing the GTM test and examining the groups which did well in the test, they identified the following characteristics of effective collaborative group work, suggesting that the GTM test might be measuring these characteristics as effective group thinking:

- encouraging each other, expressions of humility;
- giving clear elaborated explanations, equal participation with everyone in the group actively involved in each problem;
- actively seeking agreement from others, not moving on until it is clear that all in the group understand;
- asking open questions;
- sharing smiles and laughter;
- willingness to express intuitions, indicating mutual respect in tone and responses;
- taking time over solving problems seen in accepting pauses and giving elaborated explanations when asked.

2.3 Collaborative learning in mathematics

The characteristics discussed in the above sections might be seen in group work in mathematics, particularly in mathematical problem solving. In addition to utilising basic skills such as undertaking the number operations correctly, building and using representation and images of mathematical objects are also important. For example, Pirie and Kieren's model (1994) describes developmental processes for mathematical thinking and understanding. In this model the development of mathematical images and representations plays an important role. This is described as growth in understanding, with eight potential layers for mathematical understanding; Primitive knowing, Image making, Image having, Property noticing, Formalising, Observing, Structuring and Inventing. Martin and Towers (2015) further apply this model to collaborative group work. They take collaborative thinking and understanding as 'improvisational process' (Martin and Tower, 2009), and examine whether similar developmental paths are observed in collaborative problem solving. They recognise that Collective image making, Collective image having and Collective property noticing, derived from Pirie and Kieren's model, can be a useful model to describe collective mathematical thinking (Martin and Towers, 2015).

For example, suppose a group of students are solving the revolving door problem (Appendix). In order to determine the maximum number of people that can enter the building through the door in 30 minutes it might be necessary to create an appropriate image of the situation (Collective image making/having), and then examine the problem situation based on the created image (Collective property noticing). However, in mathematics learning students' problem solving processes might not be straightforward and the created image might not work to find solutions for the given problem. In this case they might have to *fold back* to the earlier stages such as Collective image making, which Pirie & Kieren (1994) considered a crucial step for conceptual growth in mathematics.

Martin and Towers (2015) conclude that 'there emerges a need to look to others within the group to participate in this process and it is through this shared action that Collective image making, Collective image having, and Collective property noticing occurs (pp. 16-7) – this echoes what Wegerif et al. (2017) found as 'characteristics of effective group thinking' in their examinations of the data from GTM listed above. In this paper we shall investigate the relationships between the thinking measured by the GTM test and mathematics problem solving.

3. Methodology

In order to answer our research question, we conducted empirical studies in Japan and the UK. In this paper we report on the data from 164 students (10-12 year old)

DOI:10.1080/03004279.2019.1701513

from 4 schools. These students took the GTM test (see 3.1) and the Maths Group tests (see 3.2), and some of the group work was video recorded (see 3.3). The following table summarises our data sets.

Table 1 Summary of data sets

School / Grade	Population (N)	Number of groups (N)	GTM test	Maths tests
School P Y7 (UK)	30	10	10 videos	10 videos
School A G6 (Japan)	37	12	3 videos	3 videos
School A G5 (Japan)	36	12	1 videos	1 videos
School B G5 (Japan)	36	12	1 videos	1 videos
School O G5 (Japan)	25	9	1 video	1 video

In what follows, we shall describe each test, the participants and the data analysis approach.

3.1 Group thinking measure test

The GTM test (Wegerif et al, 2017) consists of two tests A and B each with 15 graphical puzzles which are carefully matched for difficulty (examples are shown below). The test has been used with more than 300 school children so far (Wegerif, et al. 2017). The individual test was conducted first, with half of the group using Test A and half Test B, and the arrangement into triads was based on their individual scores. For each question, students have to choose which graphical image should fit into '?' based on the patterns and properties of the other 8. For example, the correct one the answer is '4' by, for example, seeing small circles 'outside' as addition and small circles 'inside' as subtraction horizontally (7+0=7)/vertically(1+6=7), and the no. 4 has 7 small circles outside). In addition, combining the use of this group measure with videoing and transcribing of the interaction in one to three focus groups per class of students enables a connection to be made between discourse and the group thinking outcome measure which is suggested by the extensive literature review conducted by Wegerif et al. (2017). The difference in the scores of the individuals compared to the groups will give us a measure of how well the group thinks together. If the group score is higher than any of the scores of the individuals

Education 3-13
DOI:10.1080/03004279.2019.1701513

making up the group then that indicates that the group is working well. If the group score is lower than any of the individual scores then that indicates that the group is not working well.

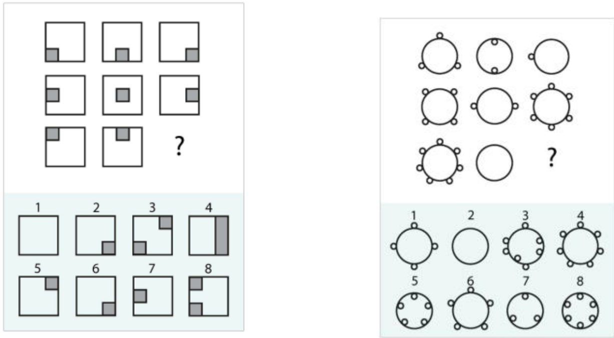


Figure 1 Test items in Group thinking measure test

3.2 Mathematics test

For the mathematics test, we used two problems – firstly, the revolving door problem taken from PISA 2012 (‘Door’), and secondly, the calculation triangles (Calc. Triangles), shown in the Appendix. These problems were chosen because students are required to make and use certain mental ‘images’, to recognise hidden patterns, and to check their solutions and answers in order to successfully complete the problems, which are also required to solve the GTM. We gave 4 points for the door problem (correct thinking and correct answers), and 30 points for the triangle problem (each triangle has three numbers to fill).

In addition, calculation triangles are examples of the *Substantial Learning Environments* (SLEs) which Wittmann (1995) proposed as a principle for mathematically rich problems. We considered these predesigned problems appropriate for our purposes, as they encourage students to work together in order to make sense of hidden patterns in the triangles. For example, in the bottom right triangle, answers can be found by formulating equations but also can be solved by using the number patterns which are common in the triangles. Finding patterns are also the key strategy for the GTM test, and we expected that if pupils performed well in the GTM test, then they might be able to use such strategy to solve Calc. Triangles.

DOI:10.1080/03004279.2019.1701513

We observed this in our pilot study with Year 6 UK pupils.

3.3 Participants and data sets

We tried to match up the participants' background in terms of their mathematical abilities through convenient sampling although we are aware that many cultural factors might affect in-test performance. In total 2 UK schools and 4 Japanese schools agreed to participate; complete data sets are presented from 1 UK and 3 Japanese schools (Schools A, B and O).

For the UK school, 30 pupils from Y7 at School P (11-12 year old; M=18, F=12), participated. Their mathematical abilities are recognised by their class teacher as the second highest group in the year group, meaning that their achievements are higher than average students in the UK school context. 37 G6 pupils (11-12 year old; M=14, F=23) from School A (a University attached school) participated and their achievements are higher than the average students, although their class teacher disclosed that some of them feel anxious about mathematics. Also, 36 G5 pupils from School A, 10-11 year old, M=16, F=20), 36 G5 pupils from School B (another University attached school, M=17, F=19), and 25 G5 pupils from School O (a Japanese state school, M=13, F=12) participated. Although they have had different learning experiences based on each country's national curricula for mathematics, both sets of students have learnt enough mathematical knowledge and skills to solve the mathematics test problems. Also none of students had experiences of GTM, Door and Calc. triangles before this research. We consider that choosing these pupils from each country under the conditions described above was a reasonable starting point for our comparison.

We also video recorded each groups' work for both GTM (group) test and maths tests for the UK pupils. Meanwhile, we just chose 3 groups for video-recording from the Japanese pupils due to limited time for data collection. For the UK sample, the first two authors undertook the data collection in November 2015-June 2016, whereas for the Japanese school the first author conducted the data collection in

March 2016 and 2018. Video-data from Japan was first transcribed and translated into English by the first author, and then checked by the second and third authors.

3.4. Data analysis

3.4.1 Quantitative data analysis

The pupils undertook Door and Calc. Triangles in groups of 3 (occasionally 4), after they solved the GTM test. We first devised quantitative data from these, i.e. descriptive/inferential statistics of GTM and group performances of the maths tests. For further analysis, we also explored how many 'Value Added groups' etc. were in each school. These groups were defined by Wegerif et al. (2017, p. 45) as follows.

- Value Added Groups (VAG): Groups that score over a standard deviation (SD) more than the highest score of any of the individuals in the group;
- Value Detracting Groups (VDG): Groups where the group result is more than one SD lower than the highest individual within the group;
- Value Neutral Groups (VNG): Groups that score between these two.

For example, assume 3 students with their GTM individual scores 8, 5 and 3 (out of 15) performed 11 out of 15 for their GTM group score. A population SD n is 2.06, and the sum of the highest individual score 8 and the SD is 10.06, but their group GTM score is actually higher than this. Therefore, this group is considered as a VAG. All calculations were done using the free statistics software *R* ver 3.3.3.

3.4.2 Qualitative data analysis

We also conducted a qualitative analysis for the video data; as Wegerif et al. (2017) stated the GTM test 'is particularly useful in integrating qualitative interpretations of group processes using videos of groups working together around the tests, with quantitative measures of the success or failure of group thinking' (p. 44).

In this paper, we have selected the following 4 groups from our data in order to illustrate pupils' problem solving processes, two each from the UK school P (Y7

DOI:10.1080/03004279.2019.1701513

pupils) and the Japanese School A (G6 pupils). We chose these groups as we noticed interesting group discussions in our initial data analysis (e.g. UK G2 is a VAG but they did not talk much, JP G2 is also a VAG but 1 male and 2 females worked almost separately). We acknowledge that these samples are limited, but they also provide rich sources of information on how the children worked in groups for the GTM and the maths tests (see section 4.2).

Table 2 Selected groups' performances in the test in the UK and Japan

Group	Students	Ind. GTM	Grp. GTM	Door	Calc. Triangles
UK G2	All males	6, 5, 6	10	1	24
UK G10	2 females & 1 male	9, 9, 9	8	1	21
JP G1	All females	14, 13, 13	15	4	27
JP G2	2 females & 1 male	7, 8, 7	12	2	24

We examined the video data from the four groups in terms of the three points of types of talk, collective image making/having, and characteristics of group work. In our qualitative analysis, we gave particular attention to the following aspects informed by our literature review:

a. what characteristics of collaborative group work were observed, informed by Wegerif, et al. (2017)
b. how pupils engaged with group work by looking at the types of talk (e.g. explorative/cumulative/disputational) informed by Mercer and Sams (2006) and Rabel and Wooldbridge (2012);
c. what kind of images they were collectively establishing in order to solve problems, informed by Martin and Towers (2014).

Although we are aware of the importance of scaffolding their group work by teaching how to engage in exploratory talk by using ground rules for talk (e.g. Alexander et al., 1992; Mercer and Sams, 2006; Rabel and Wooldbridge, 2012), we did not take this approach. We also focused on collaborative group work where teachers' or instructors' interventions are minimal. This is because our original study interests were to examine relationships between general group thinking and group mathematical thinking, measured by the tests before interventions. This clearly limits some aspects of the collaborative group work observed in the research for this paper; we will return to this issue later.

4. Findings

4.1. Overall test performance and relationships between GTM and mathematics tests

4.1.1 Differences between individual and group work in the tests

Tables 3 and 4 summarise descriptive statistics (mean scores with %s and standard deviations (SD)) in each test in the UK and Japan.

Table 3: Individual/Group results from GTM

Schools	Ind. GTM Mean (%); SD	Grp. GTM Mean (%); SD
School P Y7	9.1 (61%); 2.1	10.4 (69%); 1.3
JP School A G6	10.7 (71%); 1.9	13.4 (89%); 1.8
School A G5	10.6 (70%); 2.1	13 (87%); 1.3
JP School B G5	9.9 (66%); 2.2	13 (87%); 1.8
JP School O G5	7.9 (53%); 2.7	11.8 (79%); 1.4

Table 4 Group results from Maths tests

Schools	Door Mean (%); SD	Calc. Triangle Mean (%); SD
UK School P	1.7 (43%); 1.3	19.3 (64%); 6.5
JP School A G6	3.5 (88%); 0.9	26.8 (89%); 1.6
JP School A G6	3.3 (83%); 1.1	25.6 (85%); 3.7
JP School B G5	3.2 (80%); 1.3	27 (90%); 2.4
JP School O G5	3.3 (83%); 1.0	25.2 (84%); 2.6

As we can see in table 3, both UK and Japanese pupils benefitted by working within small groups as the group scores are higher with smaller standard deviations. Additionally, from table 4 we can see that Japanese pupils did very well in the collaborative mathematics test (Door + Calc. triangles).

DOI:10.1080/03004279.2019.1701513

However, from the GTM scores and mathematics scores, we could not find a strong direct correlation between these two scores. For example, there was a weak co-relation between the GTM and the Maths scores (Door + Calc. triangles) (0.38, $p<.05$). Also although a linear regression model 'Maths score'=1.07x'GTM score'+14.7 is significant ($p<0.01$), the adjusted R^2 is 0.12 (12% of the variance explained), indicating that this model does not adequately explain direct (linear) relationships between GTM and group mathematics scores.

4.1.2. Value Added/ Neutral/Detracting groups

Regardless of the country, it would be expected that group scores might be higher because some weaker pupils are benefiting from working with stronger ones. In order to investigate this further, we tried to identify how many Value Added/ Neutral/Detracting groups (Wegerif, et al. 2017) there were in each country, summarised in table 5 below.

Table 5 Value Added/Neutral/Detracting groups

Schools	Value Added	Value Neutral	Value Detracting
School P (UK Y7)	4 (40% of 10 groups)	3 (30%)	3 (30%)
School A (JP G6)	7 (58.3% of 12 groups)	5 (41.2%)	0 (0%)
School A (JP G5)	7 (58.3% of 12 groups)	5 (41.2%)	0 (0%)
School B (JP G5)	5 (41.7% of 12 groups)	6 (50%)	1 (8.3%)
School O (JP G5)	5 (55.6% of 9 groups)	4 (44.4%)	0 (0%)

Again we do not claim that the results represent all UK and Japanese schools. However, as we can see, about half of the groups are assessed as 'Value Added Groups' in Japan, and only 1 group as 'Value Detracting Group', whereas there were more 'Value Neutral or Detracting Groups' in the UK schools. This might again explain why Japanese pupils' performances are much better than the UK pupils' at least in our sample.

4.2. Pupils' collaborative problem solving processes

The numerical data from the GTM and the maths tests presented in the previous section suggest that whilst both UK and Japanese pupils benefitted from collaborative group work, Japanese pupils gained more from collaborative group work than the UK pupils. In order to investigate the ways the pupils worked in these tests, we examined their collaborative problem solving processes using video data derived from 4 groups (see table 2 in section 3.4.2) in terms of. 1) characteristics of group work, 2) types of talk (disruptive/cumulative/explorative), and 3) their images/representations used for problem solving

4.2.1 Characteristics of group work

Overall, our video analysis suggested that there are no straightforward patterns in pupils' group work. For example, UKG2 and JPG2 are VAG as their group scores exceeded their individual scores, but their group interactions are quite different in terms of characteristics of group work. For example, in UKG2, there was little cohesion within the group, or encouragement, and two of the three boys dominated much of the discussion except GTM Q11. JPG2 did not share physical spaces much, each one used his/her finger to point at the question or different potential answers but the others did not join in. They also gave their answers by guessing, but not stating their reasoning. They sometimes rushed to find answers rather than examining their thinking.

The UKG10 was a VDG, but in terms of the characteristics of group work, they worked well together, using humour as a strategy to cohere as a group. There were some extended explanations of reasoning, but this tended to be uninvited, rather than in response to other group members' requests for clarity or explanation. There was some rush to get through the material, little checking of answers once they had

DOI:10.1080/03004279.2019.1701513

reached a decision. The group often began by recognizing that ‘Again, this is really hard…really, really hard’.

Among the four groups, JPG1 scored 15 (100%), 4 (100%) and 27 (90%) for the GTM, Door and Calc. Triangle, respectively. Although we need to take into consideration that their individual GTM scores were also high (14, 13 and 13), they worked very efficiently as a group, showing effective group work throughout their problem solving for both the GTM and the maths tests, e.g. they encouraged each other, giving clear elaborated explanations, equal participation with everyone in the group actively involved in each problem, etc. They often smiled during their problem solving, sharing their fingers/pens on the test paper and so on.

4.2.2. Types of talk

Our analysis suggests that the students used different types of talk, but in general the UK groups’ talk was classified as disputational, and Japanese groups as explorative. For example, the students in JPG1’s talks were mostly characterised as ‘explorative’. They often used explicit reasoning to explain their ideas, asking critical questions to clarify each other’s ideas, etc. JPG2 was a VAG, which was due to the two female pupils (JPG2 F1 and F2) who retained their exploratory talk by dealing with the male pupil (JPG2 M) by either responding gently or strategically ignoring his random guesses of the answer. This strategy worked for easier questions, but failed to enable them to correctly answer more difficult questions such as B11, where they were disturbed by the male pupil’s random guess:

[Solving GTM B11, figure 1 right]

JPG2 M	Is it No. 6? No. 5, No. 5?
JPG2 F2	Maybe No. 6?
JPG2 F1	Is this…
JPG2 M	This one is 6, and there are 6 inside, so No. 6, or No. 8?
JPG2 F1	But this one is, 6, 6, 6, 666 and why there is no 6?

JPG2 M	Maybe No. 8? No. 8!
JPG2 F1	Multiplication, or addition?
JPG2 M	Maybe not No. 8.
JPG2 F2	This is, 3, 2 and 1.
JPG2 M	No. 8. Hurry!
JPG2 F2	4 and 2 is 6, and 7
JPG2 F1	Yes.
JPG2 F2	1, 2, 3, 4, 5, 6, 7
JPG2 M	Maybe No. 6.
JPG2 F1	No. 2?

Their interaction was similar in the mathematics tests – although they maintained their resilience for problem solving, their talk was rather cumulative, and the male pupil just threw his ideas in without any reasoning or explanations, resulting in a failure to solve harder questions such as the sixth triangle outside numbers 187, 188 and 189 and find insides), as follows:

JPG2 M	What? I do not understand this at all! Uh, I think, we can make 99 and 99.
JPG2 F2	Yes I agree.
JPG2 M	99 and 99?
JPG2 F1	This one? OK, what?
JPG2 M	What? It became 198…
JPG2 F2	No, that's not good, OK, uh.
JPG2 M	80…
JPG2 M	Either one could be 89.
JPG2 F2	No, that is not good like that.
JPG2 M	Is it 94?
JPG2 F2	What? Why?
JPG2 F1	OK, how about 188 is divided by 2?

They continued to try different numbers, but instead of offering possible strategies, JPG2 M kept throwing in numbers which did not lead them to a correct combination of three numbers (i.e. 93, 94 and 95).

Interestingly, although UKG2 was a value added group, their group work was in general disputational, exchanging their ideas very little, and in fact one male

DOI:10.1080/03004279.2019.1701513

student dominated their group work. However, in their GTM question A10 and A11, they exchanged their reasoning in an exploratory manner. For example:

[Solving GTM A11]

UKG2 M3	It's up then it's down, so it's going to be in the middle.
UKG2 M1	No, because it's going like that, on top there.
UKG2 M1	There, middle, there, it's going to be on the bottom one here (points at ?), so it's either that one (points at a4) or that one (a5).
UKG2 M2	I think it is number four because of the equals sign and the circle at the bottom.
UKG2 M1	Could be.
UKG2 M3	Arrows probably want to go that way (motions across page).
UKG2 M3	It could be this one (answer 1) It can't be that one (answer 2).
UKG2 M1	It can't be that one (answer 1) because that would be on the top there and we've already got it there.
UKG2 M3	Look this one, look, it is at the bottom.
…	
UKG2 M1	Look there is always two equals in.
UKG2 M1	(points at answer 4).
UKG2 M3	Look there's no equals there and there is an equals there.
UKG2 M1	No, if you look, there, equals, equals (points at middle row) so it can be equals.
UKG2 M2	Hang on, if you look at all the rows, they have two equals.
UKG2 M1	Yeah. That's why I think it will be 4.
UKG2 M2	Four.
UKG2 M3	So you think it is four?
UKG2 M2	Yup.
UKG2 M3	OK, you guys think it is (marks 4 and turns page).

Here, we can see that they worked more strategically by:

- eliminating possible wrong answers (A10 and A11, e.g. 'It cannot be that one');
- stating their reasons explicitly (A11, e.g. 'It can't be that one (answer 1) because that would be on the top there and we've already got it there', 'No, if

you look, there, equals, equals (points at middle row) so it can be equals', etc.);

- trying to establish images of the problem (A11, e.g. 'Arrows probably want to go that way (motions across page)', 'Hang on, if you look at all the rows, they have two equal')
- asking others for agreement (A11, 'OK, you guys think it is') etc.

As a result, answering correctly for these two questions might have affected UKG2' group score (10) which was better than their individual scores (6, 5, and 6). However, UKG2 only showed their successful problem solving in these two questions.

4.2.3. Use of images/representations for problems solving

JPG1's image making was effective as they tried to seek patterns in each item in the GTM, asking 'Are there any patterns in this problem?' or 'If circles inside mean addition…' etc. They quickly established effective images for the Door or Calc. triangles problems. For example, question B11(Figure 1 right) had been shown to be challenging difficult in our previous study (about 33% of the participants answered correctly in Wegerif et al, 2017), but this group solved this question well by making sense making of what patterns might be hidden in this problem:

JPG1 F2	Do you see any pattern?
JPG1 F1	It seems… like opposite, and then 4+2=6.
JPG1F2	What? Then…
JPG1F1	I am not sure… But what patterns are there?
JPG1F2	4+2… 6-2=4.
JPG1F1	Or not?... Or circles 'insides' represent addition?
JPG1 F2&3	Uhm…
JPG1 F1	I think that is related. Or, but, 4_,
JPG1 F3	Is that 8?
JPG1 F1	But 1, 2, 3, 4, 5, 6, 7, 4+3=7, and 2-2=0, then 1+7=…
JPG1 F2	7, 7!

They maintained similarly effective group work throughout the maths test, resulting in one of the best performance groups in our sample.

DOI:10.1080/03004279.2019.1701513

JPG2 struggled to make effective images for Q11-15 in GTM, Door Q2 and the sixth and tenth Calc. Triangles. However, they sometimes folded back to image making, and this saved wrong answers by mistake. For example, JPG2 answered correctly for B10 in which answer 2 (vertical lines) and 7 (horizontal lines, correct answer) look very similar:

JPG2 F1&F2	No. 2, No. 2, No. 2.
JPG2 F2	Definitely No. 2.
JPG2 M	No. 2 and No. 7.
JPG2 F1	Wait, wait!
JPG2 M	This one, horizontal lines, horizontal!
JPG2 F1	Horizontal, that was close.
JPG2 M	No. 7 because it has got horizontal lines.

The group UKG10 did not do well in their group maths test either. In general, they struggled to establish 'images' for the problem, did not engage in productive dialogues in their problem solving:

[Solving GTM B11, figure 1 right]

UK G10 F1	They're like Christmas lights.
UK G10 F2	Yeah!, like snowflakes with no pattern.
UK G10 M	Snowflakes slash Christmas lights.
UK G10 F2	4.., &3.
UK G10 F2	(points to questions in silence).
UK G10 M	2, 3, (points to top row of questions).
UK G10 F2	Right ok. (pauses) I don't get it! How does it go from that to that, but that one.
UK G10 M	(points at questions and mumbles).
UK G10 F1	I think it is that one (points to answer 5) or that one (points to answer 8). It's gone from THAT (points at q top row, last col) to THAT (points at answer 8).

As we can see above, they started describing the puzzle as 'Christmas lights' or 'snowflakes with no pattern', which were not useful to solve this question (UK G2 showed similar group work for this question).

DOI:10.1080/03004279.2019.1701513

5. Discussion

Collaborative group work is recognised as one of the important approaches in education, and there are in fact numerous studies in this area. However, what was missing was how to measure group work or group thinking by considering individual performances; this is why Wegerif et al. (2017) developed a simple test (GTM) to measure students' group thinking. In this paper, as a first attempt to apply this GTM, we explored collaborative group in mathematics around the two research questions i)'What are the relationships between general collaborative group thinking measured by the GTM and mathematical thinking measured by tests, and ii) What can we learn from collaborative problem solving process in GTM and maths activities?'.

We have collected data from Japan and the UK, and analysed data both quantitatively and qualitatively. As an answer for the first research question, the relationship between the GTM and our maths test (in our case Door and Calc. Triangles) is still not certain as suggested by the correlation and linear regression analysis (Of course this is limited to the mathematics problems used in this study). However, almost half of the Japanese groups were VAGs, and this suggests that they were good at collaborative thinking. Furthermore, the Japanese groups did well in their maths test (e.g. On average 80% or above). Therefore, examining if a group is VAG or not in their GTM scores is more useful to identify more mathematically effective groups rather than just directly comparing their group GTM scores and mathematics scores. We also consider this might provide new insight for examining findings reported in previous studies. For example, Dahl et al. (2018) reported the girls' group did better than boys although "the boys' skills in mathematics exceed those of the girls" (p. 610), but this could be explained by the GTM - the girls' group might be a VAG, and boys VNG or VDG.

Although interactions within the selected groups were very complex, it seemed that in both Japanese and UK groups, successful problem solving might be based on the use of strategies such as eliminating possible wrong answers, stating their reasons, trying to establish images of the problem, asking others for agreement etc. These are all recognised as useful strategies in related studies in both general and mathematical group work (e.g. Wegerif, et al. 2017; Martin and Towers, 2015). Also, it was evident that exploratory type talk plays important roles in their problem solving, which have again been pointed out in existing studies (e.g. Mercer and Sams, 2006; Rabel and Wooldbrige, 2012; Newman, 2017; Dahl, et al. 2018). In addition to the factors for effective group work in existing literature, we can learn that careful checking of work, and thus overcoming silly mistakes, combined with exploratory, rather than disputational types of talk appear to be related to better outcomes. These points should be carefully considered when we develop a pedagogical model for better collaborative group work. In summary, we can learn that the use of image making strategies with exploratory talk are one of the key elements of successful group work, and VAG/VNG/VDG groups suggested by GTM might be measuring this aspect of group work – this is our tentative answer for the second research question.

6. Conclusion

This paper takes a first step in exploring how the GTM can be used in education. Our paper focused on mathematics, and it was encouraging that focusing on the VAG/VNG/VDG which was suggested by GTM might be used as indicators to identify effective groups for maths group problem solving – this is one of our main conclusions of this paper. The results from the GTM can be used as indicators of effective group work in specific subject areas such as mathematics, although the subject specific thinking and strategies for problem solving should not be underestimated, e.g. encouraging students to make mathematically meaningful images or representations to solve specific mathematics problems. Of course, we only used very limited contexts of mathematics learning, and we do not intended to over-generalise our results. Therefore, it is necessary to extend our investigation by using wider examples from mathematics as well as increasing sample sizes, particularly from the UK. Also, another interesting future study would be to compare the GTM and in other subject areas, e.g. history, geography, science, religious studies with the enquiry approach, etc. Undertaking such a study will enrich our understanding of students' collaborative learning and the process of co-construction of knowledge as well as developing a better model for collaborative learning.

References

Alexander, R. J., Rose, A. J. & Woodhead, C. (1992) Curriculum organisation and classroom practice in primary schools: a discussion paper (London, DES).

Baines, E. Blacthford, P. Kutnick, P, Chowne, A, Ota, C. Berdondini, L . (2009) *Promoting Effective Group Work in the Primary Classroom: A Handbook for teachers and Practitioners.* Routledge, London and New York.

Baines, E, Blatchford, P and Webster, R. (2015) The challenges of implementing group work in primary school classrooms and including pupils with special education needs, *Education 3 -13*, 43:1, 15 -29.

Bruner, J. (1985)_Vygotsky: An historical and conceptual perspective. *Culture, communication, and cognition: Vygotskian perspectives* , 21-34. London: Cambridge University Press.

Dahl, H., Klemp, T., & Nilssen, V. (2018). Collaborative talk in mathematics–contrasting examples from third graders. *Education 3-13,* 46(5), 599-611.

Dawes, L. (2001) 'Interthinking - the power of productive talk.' In Goodwin, P (ed) *The Articulate Classroom: Talking and Learning in the Primary School.* Routledge: London and New York.

Dawes, L. Mercer, N and Wegerif, R. (2004) *Thinking together: A programme of activities for developing speaking and listening and thinking skills for children aged 8 – 11.* The Questions Publishing Company, Birmingham.

Dewey, J. (1916) *Democracy and Education*. Republished 2008 Wilder Publications, Radford Virginia.

Francisco, J. M. (2013). Learning in collaborative settings: Students building on each other's ideas to promote their mathematical understanding, *Educational Studies in Mathematics,* 82(3), 417-438.

Gillies, R. and Ashman, A.(eds) (2003) *Co-operative Learning. The social and intellectual outcomes of learning in groups*. Routledge Falmer, London.

Hertz-Lazarowitz and Zelniker, T. (1995). Cooperative Learning in the Israeli Context: Historical, educational and cultural perspectives. *International Journal of Educational Research,* 23, 267 – 285.

Howe, C. (2010). Peer dialogue and cognitive development: two-way relationship? In Littleton, K. and Howe, C. (eds). *Educational Dialogues: Understanding and Promoting Productive Interaction.* London: Routledge.

Hino, K. (2015). Comparing multiple solutions in the structured problem solving: Deconstructing Japanese lessons from learner's perspective. *Educational Studies in Mathematics*, 90(2), 121-141.

Kazak, S., Wegerif, R., & Fujita, T. (2015a). Combining scaffolding for content and scaffolding for dialogue to support conceptual breakthroughs in understanding probability. *ZDM*, 47(7), 1269-1283.

Kazak, S., Wegerif, R., & Fujita, T. (2015b). The importance of dialogic processes to conceptual development in mathematics. *Educational Studies in Mathematics*, 90(2), 105-120.

Kershner, R. Warwick, P., Mercer, N. Staarman, J (2014) Primary children's management of themselves and others in collaborative group work; sometimes it takes patience…' *Education 3-13*. 42:2, 201 -216.

Martin, L. C., & Towers, J. (2015). Growing mathematical understanding through collective image making, collective image having, and collective property noticing. *Educational Studies in Mathematics*, 88(1), 3-18.

Mercer, N**.,** Dawes, R., Wegerif, R., & Sams, C. (2004) Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal*, 30, 3, 367-385.

Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language and Education*, 20(6), 507-528.

Monaghan, F. (2005). 'Don't think in your head, think aloud': ICT and exploratory talk in the primary school mathematics classroom. *Research in Mathematics Education*, 7(1), 83-100.

Newman, R. (2017) Engaging talk: one teacher's scaffolding of collaborative talk, *Language and Education*, 31:2, 130-151.

Panitz, T. (n.d.). Collaborative versus cooperative learning: a comparison of the two concepts which will help us understand the underlying nature of interactive learning. Retrieved August 24 2017, from http://files.eric.ed.gov/fulltext/ED448443.pdf

Pirie, S., & Kieren, T. (1994). Growth in mathematical understanding: How can we characterise it and how can we represent it?. *Educational studies in Mathematics*, 26(2-3), 165-190.

Pifarré, M., & Li, L. (2018). Characterizing and unpacking learning to learn together skills in a wiki project in primary education. *Thinking Skills and Creativity*, 29, 45-58.

Pifarré, M. (2019). Using interactive technologies to promote a dialogic space for creating collaboratively: A study in secondary education. *Thinking Skills and Creativity*, 32, 1-16.

OECD (2016), *PISA 2015 Results (Volume 1): Excellence and Equity in Education*, PISA, OCED Publishing, Paris, http://dx.doi.org/10.1787/9789264266490-en

Rousseau, J. (1762) *The Social Contract.* Translated by Cole, G. (1782)

Shahamamt, A and Mede, E. (2016) Integration of collaborative learning in Grade K-5 EFL classrooms. *Education 3-13* Vol 44: 6, p. 682-697.

Slavin, R. Hurley, E. and Chamberlain, A. (1995) Cooperative learning and Achievement: Theory and Research. 2nd Edition. Boston MA: Alyn and Bacon

Rabel, S., & Wooldridge, I. (2013). Exploratory talk in mathematics: what are the benefits? *Education 3-13*, 41(1), 15-22.

Education 3-13
DOI:10.1080/03004279.2019.1701513

Wegerif, R. and Dawes, L. (2004) *Thinking and Learning with ICT. Raising achievement in primary classrooms*. London: RoutledgeFalmer.

Wegerif, R., Fujita, T., Doney, J., Perez Linares, J. Richards, A. and van Rhyn, C (2017), Developing and trialing a measure of group thinking, *Learning and Instruction*. 48, 40-50.

Wittmann, E. C. (1995). Mathematics education as a 'design science'. *Educational studies in Mathematics*, 29(4), 355-374.

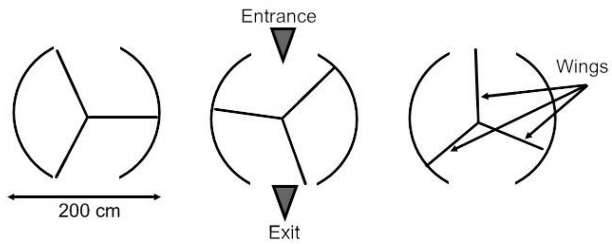
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Appendix

Door

Revolving Door

A revolving door includes three wings which rotate within a circular-shaped space. The inside diameter of this space is 2 metres (200 centimetres). The three door wings divide the space into three equal sections. The plan below shows the door wings in three different positions viewed from the top.



Question 1

a) What is the size in degrees of the angle formed by two door wings?

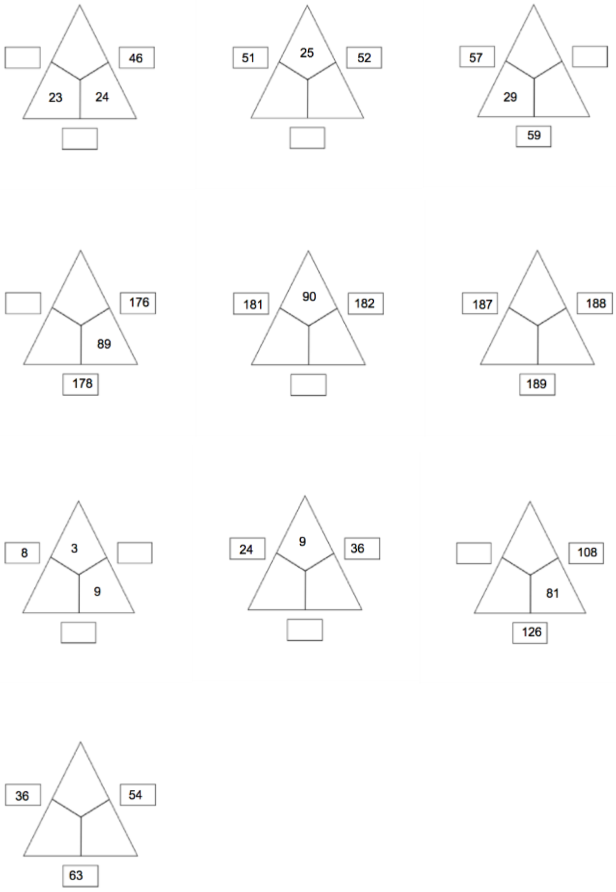
b) The door makes 4 complete rotations in a minute. There is room for a maximum of two people in each of the three door sectors. What is the maximum number of people that can enter the building through the door in 30 minutes?

Choose your answer from: A 60 B 180 C 240 D 720

The question is reproduced from https://www.oecd.org/pisa/pisaproducts/pisa2012-2006-rel-items-maths-ENG.pdf

Calc. Triangles

Complete the following number triangle in accordance with the rule you noticed.


46
23
24
51
25
52
57
29
59
176
89
178
181
90
182
187
188
189
8
3
9
24
9
36
108
81
126
36
54
63

DOI:10.1080/03004279.2019.1701513