



Group Size and Peer Learning: Peer Discussions in Different Group Size Influence Learning in a Biology Exercise Performed on a Tablet With Stylus

Jean-Baptiste Corrége and Nicolas Michinov*

Laboratory of Psychology: Cognition, Behavior and Communication (LP3C), Department of Psychology, Univ Rennes, Rennes, France

OPEN ACCESS

Edited by:

Pei Sun,
Tsinghua University, China

Reviewed by:

Heng Luo,
Central China Normal University,
China
Gerard Rayner,
Swinburne University of Technology,
Australia

*Correspondence:

Nicolas Michinov
nicolas.michinov@univ-rennes2.fr

Specialty section:

This article was submitted to
Educational Psychology,
a section of the journal
Frontiers in Education

Received: 30 June 2021

Accepted: 22 October 2021

Published: 10 November 2021

Citation:

Corrége J-B and Michinov N (2021)
Group Size and Peer Learning: Peer
Discussions in Different Group Size
Influence Learning in a Biology
Exercise Performed on a Tablet
With Stylus.
Front. Educ. 6:733663.
doi: 10.3389/feduc.2021.733663

Determining the optimal discussion group size to improve performance and learning has created an intense debate in psychology and provided mixed findings in laboratory and field settings. In a quasi-experimental study in the education field, we examined the effect of discussion group size on individual learning in a biology exercise performed on a tablet with stylus. The sample involved 102 secondary school students divided into four classes, each corresponding to one of the four experimental conditions (alone without peer discussion vs discussion in dyads vs triads vs four-member groups). They were asked to draw individually a functional schema of the human respiratory system, once before and once after discussing with peers (or reflecting alone). Both drawings were evaluated by four independent coders, and the learning gain for each student was computed from these evaluations. Results revealed that learning gain was greater for students discussing in four-member groups than for those in the other conditions. Additional analyses suggested that this effect was moderated by the students' gender, with the learning gain being greater after discussion in four-member groups among females only. These findings suggest that group size of four individuals might be the optimal configuration to improve peer learning.

Keywords: biology, gender, group discussion, group size, peer learning, tablets

INTRODUCTION

Working in groups is a frequent practice in the K-12 science classroom. In this context, determining the number of students who should work in a group is problematic. As some groups often have more members than others, an ideal group size is difficult to determine in learning settings. Identifying the optimal group size in peer discussion and its impact on learning is a crucial research issue that can have a determining influence on students learning. It may also provide some useful recommendations for teachers to improve group learning in their classroom. Based on the literature about the effects of group size on different outcomes (performance and learning) and the mixed findings generally observed, there is an interest to extend research to complex tasks or exercises in secondary education settings among students using technology-mediated learning environments. The aim of the present study was to examine the effect of discussion group size on learning gain in a biology exercise performed on a tablet with stylus.

Effects of Group Size on Performance

Based on a social-cognitive approach in the field of social psychology, the pioneer research considering the effects of group size on performance has mainly been carried out in the laboratory and organizational settings. It has led to contradictory findings on a set of dependent variables such as performance, distribution of participation, conformity and satisfaction (Lorge and Solomon, 1959; Thomas and Fink, 1963; Seta et al., 1976; Hill, 1982). Since, there is an intense debate about the minimum number of people in a “group,” with some researchers considering two individuals working together in a dyad as the smallest group size (Williams, 2010), while others state that a group is composed of three (i.e. triad) or more people (Moreland, 2010; Tasca, 2020). The term “group size” itself remains a vaguely defined concept with in some studies small groups consist of four or more members (Wheelan, 2009), whereas in others they are limited to two or three members (Yetton and Bottger, 1983).

Studies from social psychology suggest that individuals working in dyads perform better than those in triads or larger groups of four or more people (Levine and Moreland, 2004), or individuals working alone (Taylor and Faust, 1952; Schultze et al., 2012). A large number of studies have demonstrated that increasing the number of participants in a group may reduce the individual motivation and effort to work collaboratively on a task, demonstrating a “social loafing” effect in large groups (Ingham et al., 1974; Petty et al., 1977; Latan  et al., 1979; Karau and Williams, 1993). For instance, in a study in which participants had to solve intellectual problems of various difficulty, either as individuals or in same-sex groups of two, three, six, or ten members, it was shown that group performance decreased as group size and task difficulty increased (Bray et al., 1978). In the same vein, Wheelan (2009) found that work groups of three to eight members operating in organizations were significantly more productive than groups with nine members or more. In a study involving 87 groups of two to six members performing a collaborative task, Yetton and Bottger (1983) found that performance did not improve for groups larger than four. Similarly, numerous studies working on social dilemma tasks found that cooperation decreased with increasing group size (Hamburger et al., 1975; Komorita and Lapworth, 1982). For instance, in comparing three- and seven-member groups, Hamburger et al. (1975) showed that smaller groups were more cooperative than larger ones.

On the contrary, other findings in pioneering research have demonstrated that larger group size may improve group performance in a wide variety of tasks (Taylor and Faust, 1952; Seta et al., 1976; Littlepage, 1991; Littlepage and Silbiger, 1992). In comparing performance between four-member groups and dyads, Seta et al. (1976) showed that groups performed better than dyads on a memorization task only under a cooperative condition, but not under a competitive condition. Using an experimental task consisting in naming an object through a series of questions, Taylor and Faust (1952) also found that four-member groups performed better than dyads, i.e. failed less and spent less time per problem, and that dyads performed better than individuals working alone on the same criteria. In their study, Littlepage and Silbiger (1992) assigned

students to groups of one, two, five, or ten participants to answer multiple choice questions, and found that group performance rose in line with increasing group size.

Similar findings were observed in a series of studies in which individuals and groups of two, three, four, or five people had to solve highly intellectual problems (letters-to-numbers problems). Results revealed that groups of more than three performed better than dyads, but also better than the best person of an equivalent number of individuals in ‘nominal’ groups (Laughlin and Ellis, 1986; Laughlin et al., 2006). One of the ways to resolve these contradictory results about the impact of group size on performance has involved considering the type of task the groups are working on (Steiner, 1972). It appears that when a solution is offered by a group member and is easily recognized as being correct (high demonstrability) the group outperforms the best performing individual (Laughlin and Ellis, 1986; Laughlin et al., 2002). Recent findings have demonstrated that increasing the group size contributes to decreasing performance for low demonstrability problems, i.e., problems for which group members fail to recognize the correct solutions proposed by others during a discussion (Amir et al., 2018).

Effects of Group Size on Learning

Another study on collaborative learning examined the effects of group size on learning in real education settings in which a crucial role is given to peer discussions and social interactions within groups. Although student attitudes toward group discussions are often negative (Clinton and Kelly, 2020), there are valuable active learning methods for engaging students in fruitful peer discussions (Prince, 2004; Smith et al., 2009; Topping et al., 2017). In active learning methods based on peer discussion such as Peer Instruction (Mazur, 1997), students are invited to discuss with their peers to improve their learning after a first answer to a question, and to give their answer again after the discussion (Vickrey et al., 2015; Balta et al., 2017; Knight and Brame, 2018; Schell and Butler, 2018). In one of the rare studies on Peer Instruction in which group size varied (two vs three vs four-member), Relling and Giuliadori (2015) did not find any significant effect on the change in answers after peer discussion in a veterinary physiology course. To our knowledge, except in this study, the number of students involved in peer discussion has not been systematically controlled in Peer Instruction and this number generally ranges from two to an undetermined number of neighbors in the classroom. Peer discussions do sometimes take place within more or less structured small groups (Michinov et al., 2015; Morice et al., 2015), but as stated by Morice et al. (2015), “group size should systematically be controlled when peer instruction takes place [...] future studies should rigorously control group size from two, including proximate neighbors in the classroom, to four members” (p. 730).

Beyond the Peer Instruction method, there is a lack of consensus about the effects of peer discussion on learning, and the issue about the optimal group size remains largely open to date in this field (Peltokorpi and Niemi, 2019). Indeed, some research suggests that small groups function better than larger groups because their members cannot attribute the responsibility

of the discussion to others (Webb, 1982, 1989), or may lack the ability to evaluate potential solutions to a problem (Schultz, 1989). On the contrary, other studies claim that increasing the number of students in a group might improve collaborative learning, leading students to benefit greatly from peer discussions due to a wider range of views (Needham, 1987).

Some research suggests that dyads are better than groups of three or more members (Slavin, 1987; Webb, 1989; Lohman and Finkelstein, 2000; Kim et al., 2020), or individuals working alone (McDonald et al., 1985; Richey et al., 2018; Kim et al., 2020). Other research recommends groups of three to four members to improve student achievement (Lou et al., 2001; Caulfield and Caroline, 2006), while other studies have tried to distinguish the differences between triads and four-member groups (Egerbladh, 1976; Wiley and Jensen, 2006). In their study, Wiley and Jensen (2006) demonstrated that triads outperform dyads, individuals working alone and the best individual in 'nominal' groups on an arithmetic problem-solving task. Similarly, Egerbladh (1976) demonstrated that triads performed better than dyads and individuals working alone, and dyads outperformed individuals working alone. A four-member group has also been proposed as the optimal size to improve performance and learning (Alexopoulou and Driver, 1997; Shimazoe and Aldrich, 2010). In comparing dyads and four-member groups before and after a discussion in a physics course, Alexopoulou and Driver (1997) found that groups of four students functioned better than dyads in terms of both group discussion processes and learning, probably because fruitful discussions are relatively constrained in dyads. In the same vein, Kagan (1992) pointed out that group sizes of four to five are best for small group learning, and Shimazoe and Aldrich (2010) reported that the ceiling on group size should be four because beyond this number the tendency to "loaf" increases with group size. Recent findings confirmed this view, revealing that performance per individual decreased as group size increased (Peltokorpi and Niemi, 2019).

However, these findings are moderated by the type of classroom setting in which students are learning (advanced vs mainstream). Using a pretest–posttest quasi-experimental design, Apedoe et al. (2012) examined the effects of group size (dyads, triads and four-member groups) on student learning in chemistry in two types of classroom setting, advanced or mainstream. They found that in mainstream classrooms, students in triads and four-member groups performed slightly better than students working in dyads, while students in advanced classrooms performed better in dyads. In their study, group sizes of three and four did not differ from each other either in advanced or mainstream settings.

Effects of Group Size in Technology-Mediated Learning Environments

A body of studies has examined the effects of group size when individuals used various information and communication technologies for online discussions with their peers and to perform richer learning tasks. Again, the results are controversial.

In a study using a group-based mobile learning environment, Melero et al. (2015) concluded that group size did not affect individual performance. Nevertheless, the results of this study revealed that students belonging to four-member groups expressed higher levels of engagement in the task. Similarly, studies in Computer-Supported Collaborative Learning (CSCL) suggested that smaller groups (three or four members) produced better performance than large groups, i.e., groups with more than five members (Strijbos et al., 2004; Schellens and Valcke, 2006). Examining group discussions on online forums, Shaw (2013) observed that small groups had higher participation rates, which indirectly influenced learning scores. A recent meta-analysis examining the effects of Computer-Based Scaffolding (CBS) on learning among students having to solve problems showed that the size effects were higher when students worked in dyads than in triads, small groups or individually (Kim et al., 2020). Another meta-analysis on mobile-Computer-Supported Collaborative Learning (mCSCL) revealed that four-member groups had better outcomes than dyads or triads (Sung et al., 2017). In using social networks for peer discussions, Sugai et al. (2019) also showed that a four-member group was the optimal size for collaborative argumentation for educational purposes.

Taken together, these studies reveal mixed findings about the effect of group size when students discuss with peers using online technologies. They also suggest the need to conduct additional studies to examine the impact of group size on learning using interactive technologies during in-person classes. Indeed, many studies have required students to discuss online, but fewer have used technologies to support peer discussions when performing exercises during in-person classes. In the present study, during in-person classes students in different sized groups were asked to perform a biology exercise with the support of oral discussion using an interactive learning environment. Unlike previous studies, online technologies for discussion were not used.

Overview of the Present Study and Hypothesis

As previous research has not provided a consensus about the optimal group size both for individual performance and student learning, we conducted a pretest–posttest study among students in a secondary school. It involved comparing students working alone to those discussing in dyads, triads or four-member groups to evaluate the learning gain between their first and second answer to a question. More specifically, students performed a biology exercise twice, once before and once after peer discussion (or individual reflection when students worked alone). The exercise required drawing a functional schematic view of the human respiratory system on a tablet with a stylus. After their first completion, the teacher displayed collective feedback on a central screen with a video projector, providing students with an overview of all their drawings in a 'thumbnail' format. These "thumbnails" did not allow a perfect visualization of all the students' work, but provided an overview of the different drawings to support their individual thinking or group discussion. In order to evaluate the learning gain, all the individual drawings were blind coded by several coders.

As students were in a mainstream classroom setting (Apedoe et al., 2012), and based on previous studies (Alexopoulou and Driver, 1997; Lou et al., 2001; Caulfield and Caroline, 2006; Wiley and Jensen, 2006), students in three and four-member groups should perform better than students working alone or in dyads. However, several studies (Yetton and Bottger, 1983; Shimazoe and Aldrich, 2010), including those using technology-mediated learning environments (Sung et al., 2017; Sugai et al., 2019), suggest that this hypothesis can be refined according to the group size of three or four members. Thus, we expected that four-member groups would perform better than three-member groups, dyads or students working alone.

MATERIALS AND METHODS

Participants

The study was conducted on 102 secondary school students (49 males, 48.04% of the sample) ranging from 11 to 13 years old ($M = 11.9$, $SD = 0.39$), all in seventh grade. They were in four classes with the same female biology teacher. As minors were involved in the study, an informed consent from the parents of each pupil was obtained. All procedures in these studies were in accordance with the ethical standards of institutional and/or national research committees for studies involving human participants and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Materials and Instruments

Tablet and stylus: The teacher and students had a tablet (Dell Latitude 5285 2.70 GHz with a 12.3-inch screen LCD 2736 × 1824) and a pressure-sensitive stylus (Dell Active Pen) enabling straight lines to be drawn with high precision (3 mm).

Interactive learning environment: The Kassis software developed by the IntuiDoc team of the IRISA Laboratory in collaboration with Learn and Go (<http://kassis-apps.com/en>) was installed on each tablet. This software allows the teacher to create lessons and share exercises with his/her students. It contains a set of features that allows students to take notes on slides, collaborate with peers on a shared whiteboard, create questions using graphics and produce drawings or sketches (see Michinov et al., 2020 for more details). In this study, only the latter feature was used to perform a biology exercise.

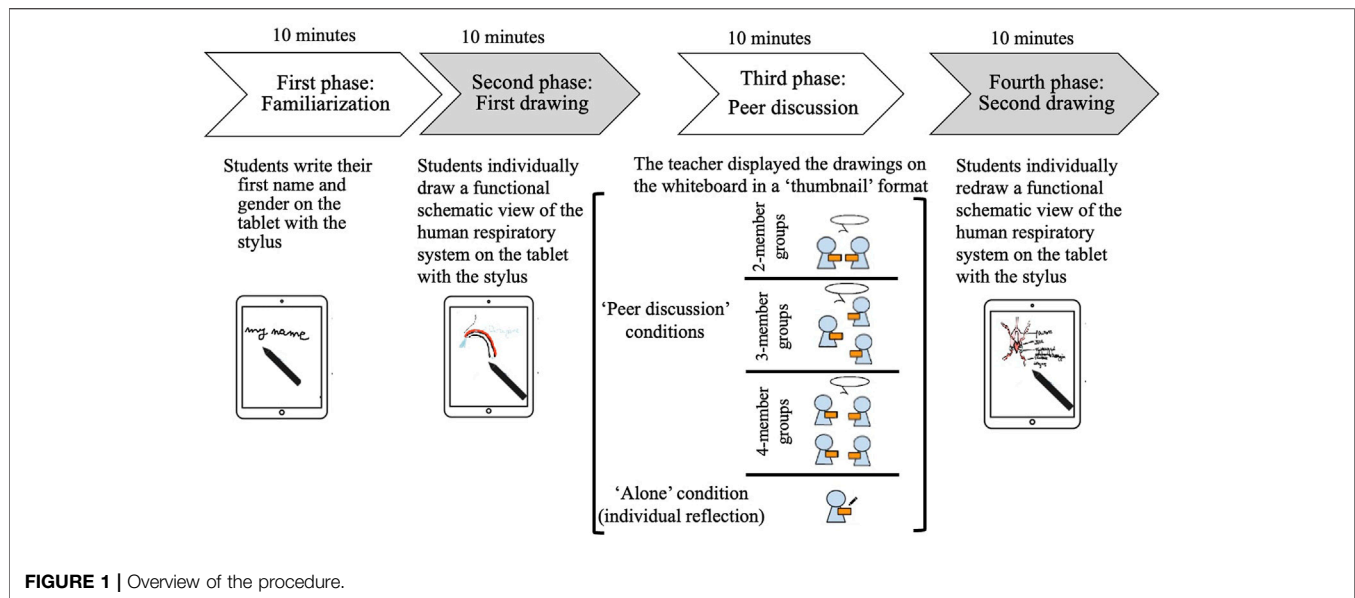
Procedure

The study took place during a biology course taught by a female teacher in four different class groups during the first semester of the year. It was presented to the students as a test of a new interactive learning environment requiring the use of a tablet and stylus to perform a biology exercise during a class. In each class, two experimenters were in the room to help students use the application and manage any potential technical problems. Each of the classes was assigned to one of the four experimental conditions involving different group sizes to which students were randomly assigned: Alone ($n = 26$, 13 females and 13 males), dyad ($n = 22$; six dyads composed of one male and one female, three dyads composed of two males,

and two dyads composed of two females), triad ($n = 30$; nine triads composed of one male and two females, one triad composed of two males and one female), and four-member group ($n = 24$; two groups composed of three males and one female, one group composed of one male and three females, and three groups composed of two males and two females). The classes were matched with an experimental condition according to the number of students in each class (e.g., a class of 30 students was chosen for allocation in 10 triads and a class of 22 students was chosen to constitute 11 dyads). As the size of the groups varied, the spatial organization of the classroom also varied across conditions to facilitate (or limit) peer discussions. In each condition, students were seated to ensure they had a clear view of the collective feedback displayed on the screen in front of them (see **Appendix A**).

The course began with the students being seated at a designated place so that groups would be randomly formed from the start of the experiment. Once seated, the teacher informed students they were not allowed to move to other groups in the classroom during the lesson. The teacher also gave each student their username to log onto the application. It consisted of the students' initials and date of birth. The teacher then introduced the experimenters and explained that, while this particular lesson used a tablet and stylus, it would nonetheless be a normal lesson. One of the experimenters then showed the students how to use the interactive learning environment.

The study itself was divided into four phases that are described in **Figure 1**. As a normal class takes 50 minutes, each phase was set to last about 10 minutes, leaving time for the teacher to answer any questions raised by the students during the remaining 10 min. The first phase consisted in a familiarization task to allow students to learn how to use the tablet and stylus and make sure they understood how it worked. The second phase consisted in producing an initial drawing which involved students individually drawing a functional schematic view of the human respiratory system on the tablet with the stylus. This exercise was chosen because students had never completed it before and they had only received some introductory elements about how the human respiratory system worked 1 week previously. The exercise was sufficiently complex for students not to be able to produce a perfectly correct schema on their first drawing and, consequently, they could improve their production from the first to the second attempt. Once they had completed their drawing, they could send it to the teacher by clicking on the "send" button and wait before proceeding to the next phase. This first version was automatically saved to be used as the pre-test. The following phase began when all the students' drawings had been received. In the third phase, the teacher displayed a collective feedback to elicit peer discussion in groups of different sizes. The feedback contained all the students' drawings on the whiteboard in a "thumbnail" format in such a way that none of them were made salient (see **Appendix B**). In the "Alone" condition, students could only examine the combined students' production displayed on the collective



feedback and were not allowed to discuss anything with their peers. Finally, in the fourth phase, students were asked to redraw individually the schematic view of the human respiratory system on a new blank page of the interactive learning environment on their tablet with a stylus without communicating with anyone. During this period, they could no longer see their peers' drawings, or their initial drawing. Their second version was recorded and served as the post-test. At the end of the lesson, students were asked to answer a single question on the interactive environment to check the efficacy of the experimental manipulation.

Measures

Manipulation check: In order to verify the efficacy of the experimental manipulation concerning the perception of the group size, students were asked how many people they had discussed with during the exercise, from zero to three.

Learning gain: This was calculated from the mean score differences between the second and the first drawings. The seven piles "sort-resort" technique was used to evaluate a score of overall quality of the drawings by four independent coders blinded to the hypothesis. This technique adapted from M. E. Shaw (1963) by Hackman et al. (1967), has been used in many studies to assess individual or collective production such as products, ideas, drawings, etc. (Craig and Kelly, 1999; Michinov et al., 2004). The coders had basic knowledge in biology, but they did not teach in this discipline. Each coder received a model of the correct schema from the teacher, as well as a folder containing color copies of all the drawings presented in a random order (pre- and post-test). They did not know the purpose of the experiment, the experimental condition to which participants had been allocated, or whether the drawing was the first or second. They first sorted the drawings into three piles (high, medium and low) on the overall quality of the drawing. Once sorted on

this criterion, each pile was then re-sorted. The "high" pile was sorted into two piles (high-high and high-low), the "medium" pile was sorted into three piles (medium-high, medium and medium-low), and the "low" pile was sorted into two piles (low-high and low-low). This resulted in seven piles corresponding to a seven-point overall quality scale ranging from 1 (low) to 7 (high). The same "sort-resort" technique was applied to evaluate whether the graphical representation of the respiratory system was adequate (e.g., whether organs such as lungs were present or not), and whether there was adequate use of legends. The evaluations on the three criteria were inter-correlated ($r = 0.60$, $r = 65$, and $r = 0.63$, all p -values < 0.001). As the inter-coder reliability was satisfactory for all the evaluations (see **Table 1**), a single composite score based on the mean scores of the four coders for the first and the second drawings was used to measure learning gain (difference between post- and pretest).

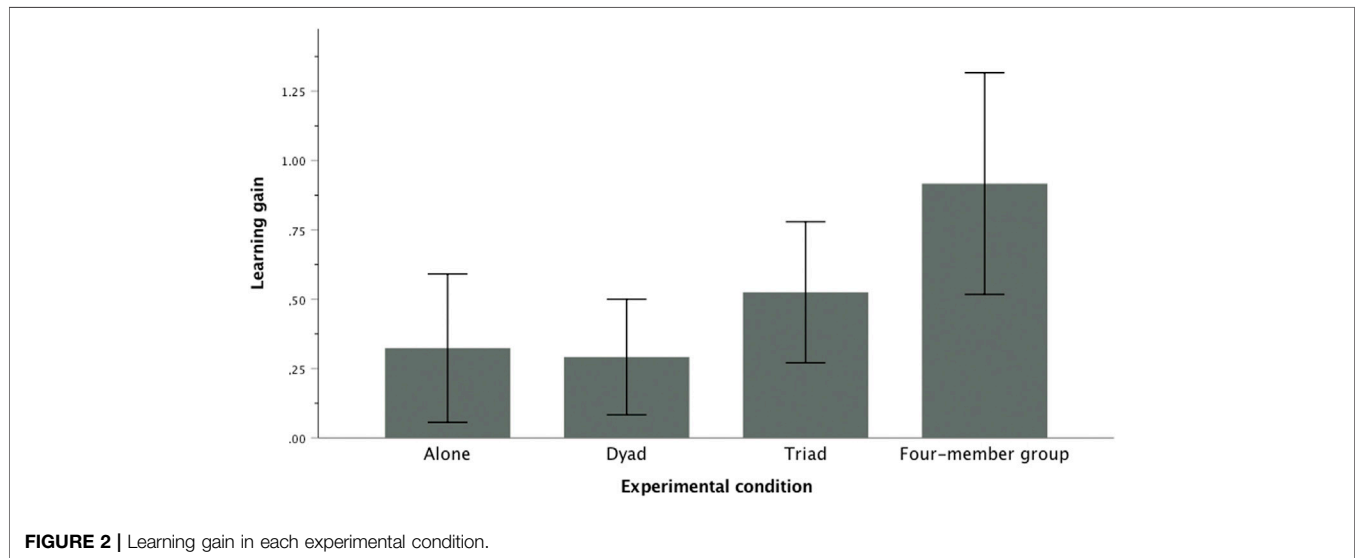
RESULTS

Manipulation Check

A Chi-square test revealed a significant difference between the four conditions, $\chi^2(9, N = 101) = 111.00$, $p = 0.001$, $\phi = 0.0605$. In the "Alone" condition, 57.7% of students reported having discussed with zero people (this relatively "low" percentage may be explained by the fact that students mentioned that they had talked to others in the classroom, but not during the biology exercise itself); in "Dyad," 68.2% indicated that they had discussed with one person; in "Triad," 76.7% indicated that they had discussed with two people; in the "four-member group" condition, 73.9% indicated that they had discussed with more than three people. These analyses suggest that students had a relatively good perception of the group size condition in which they had been placed.

TABLE 1 | ICCs between the four coders using the “sort-resort” technique on three criteria for the first and second drawings.

	Overall quality of drawings	Adequate graphical representation	Adequate use of legends
First drawings	0.86	0.81	0.89
Second drawings	0.85	0.83	0.88

**FIGURE 2** | Learning gain in each experimental condition.

Learning Gain

Preliminary analyses were computed to verify whether there were any differences between the experimental conditions on the first drawing. No difference was found between the experimental conditions on the pretest scores, $F(3, 98) = 1.54$, $p = 0.21$, $\eta^2 = 0.04$ (“Alone”: $M = 1.89$ and $SD = 0.95$; “Dyad”: $M = 2.24$ and $SD = 1.03$; “Triad”: $M = 1.78$ and $SD = 0.92$; “four-member group”: $M = 2.27$ and $SD = 1.14$). None of the Tukey’s Post-Hoc comparisons tests were significant.

An analysis of variance (ANOVA) with group size as between-subjects factor revealed a significant effect on learning gain, $F(3, 98) = 3.90$, $p = 0.011$, $\eta^2 = 0.11$ (see **Figure 2**). A significant planned comparison (1 1 1–3) revealed that students discussing in four-member groups outperformed those in all the other conditions, $t(98) = 3.223$, $p = 0.002$. Additional Tukey’s Post-Hoc comparisons tests showed that the only significant differences were between the “four-member group” condition and two other conditions (dyads and working alone). Specifically, it appeared that students involved in four-member groups ($M = 0.92$, $SD = 0.95$) improved their learning more than students working in dyads ($M = 0.29$, $SD = 0.47$), $t(98) = -2.975$, $p = 0.02$, or alone ($M = 0.32$, $SD = 0.66$), $t(98) = -2.943$, $p = 0.021$, but they did not significantly differ from those working in triads ($M = 0.52$, $SD = 0.68$), $t(98) = -2.009$, $p = 0.192$. No other differences were significant between the experimental conditions.

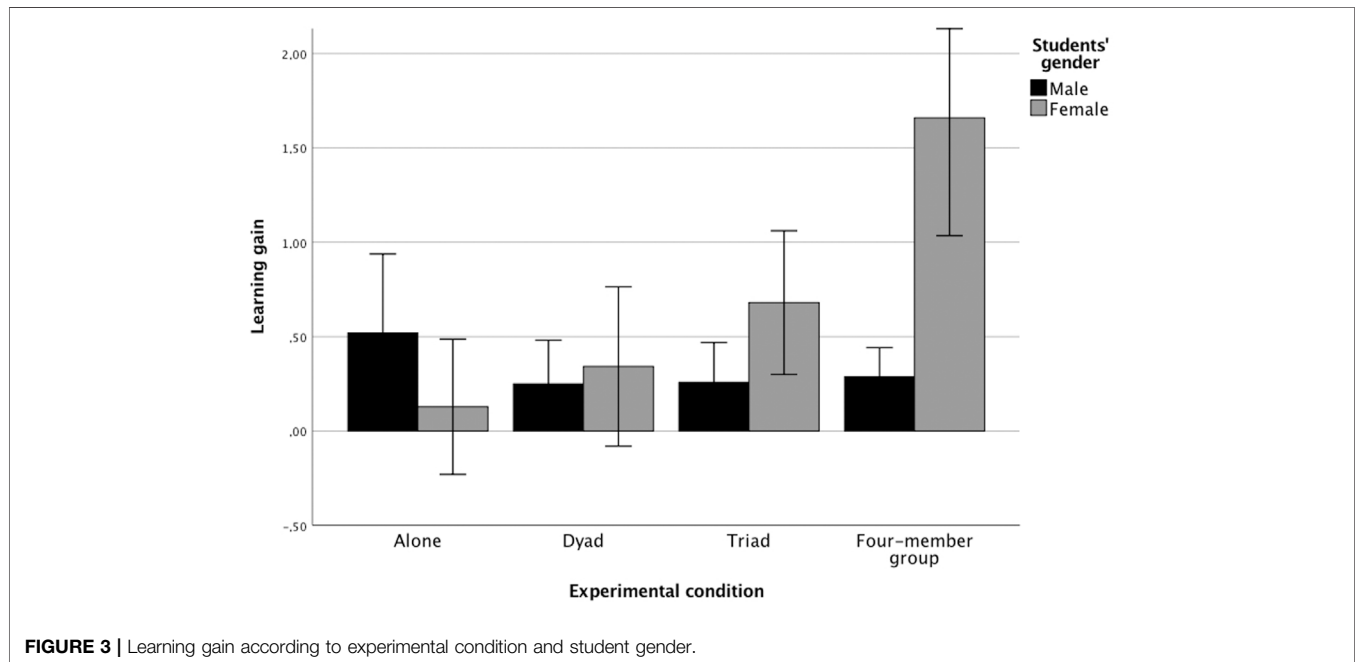
A priori contrasts yielded only a significant difference on learning gain between students discussing in four-member groups and those in all the other conditions, the results confirmed the hypothesis that students discussing in four-member groups would perform better than those in other discussion groups or those working alone.

Additional Analyses

Although not systematically controlled in the present study, the students’ gender was taken into consideration in further analyses, and treated as a moderator of the effect of group size on learning gain. A between-subject factor ANOVA with students’ gender and group size as predictors yielded that both group size, $F(3, 94) = 6.13$, $p < 0.001$, $\eta^2 = 0.12$, and students’ gender had an effect on learning gain, $F(1, 94) = 8.94$, $p = 0.004$, $\eta^2 = 0.06$. Pairwise comparison showed that female students ($M = 0.68$, $SD = 0.91$) improved to a greater extent than male students ($M = 0.33$, $SD = 0.45$, $p = 0.004$, $d = 0.01$). These effects were qualified by a significant interaction between group size and students’ gender, $F(3, 94) = 8.81$, $p < 0.001$, $\eta^2 = 0.18$ (see **Table 2** and **Figure 3**). Post-hoc comparisons showed that female students in four-member groups ($M = 1.66$; $SD = 0.93$) improved more than females who worked alone ($M = 0.13$, $SD = 0.59$, $p < 0.001$, $d = 0.03$), in dyads ($M = 0.34$, $SD = 0.59$, $p < 0.001$, $d = 0.02$), or triads ($M = 0.68$, $SD = 0.79$, $p = 0.002$, $d = 0.02$). Only a single difference explained the interaction effect for which female students improved

TABLE 2 | Means (M) and Standard Deviation (SD) of learning gain for each experimental condition according to students' gender.

	Alone			Dyad			Triad			Four-member group		
	n	M	SD	n	M	SD	n	M	SD	n	M	SD
Female	13	0.13	0.59	10	0.34	0.59	19	0.68	0.79	11	1.66	0.93
Male	13	0.52	0.69	12	0.25	0.36	11	0.26	0.32	13	0.29	0.25
Total	26	0.32	0.66	22	0.29	0.47	30	0.52	0.68	24	0.92	0.95

**FIGURE 3** | Learning gain according to experimental condition and student gender.

more than male students in four-member groups ($M = 0.29$, $SD = 0.25$, $p < 0.001$, $d = 0.03$). There was no difference between the experimental conditions among male students, although they tended to have better learning gain when alone.

DISCUSSION

The aim of the present study was to determine the impact of peer discussion on learning in groups of different sizes and potentially to try to determine the optimal size of a group in peer learning. A quasi-experimental field study was conducted in a secondary school among students following a biology course using a technology-mediated learning environment. During a learning session, they had to individually produce a schematic view of the human respiratory system on a tablet with stylus. This production was performed twice, once before and once after discussing in groups (or reflecting on their own in the control condition).

The results verified our hypothesis and thus are consistent with studies demonstrating that students discussing in four-member groups learn better than those in triads, dyads, or individuals working alone (Yetton and Bottger, 1983; Shimazoe and Aldrich, 2010; Sung et al., 2017; Sugai et al.,

2019). Based on findings found in a mainstream classroom (Apedoe et al., 2012), it could also be expected that students would perform better when they collaborated both in three- and four-member groups than in dyads or working alone. However, no difference on learning gain was found between students discussing in triads and those in the other conditions (including the four-member group condition). Taken together, the present results yielded an effect of group size on learning gain revealing that discussing in four-member groups is beneficial to learning improvement. At least two different interpretations may be proposed to explain these results.

First, it is possible that the number of interactions between students in 'large' groups may explain the benefits. Discussion is more constrained in dyads than in four-member groups (Alexopoulou and Driver, 1997), it is therefore reasonable to consider that the greater number of interactions in "large" groups than in "small" groups partly explained learning gain. However, as peer discussions were not measured in the natural classroom setting where this study took place, it would be interesting in future research to capture student interactions in groups of different size in real-time. Such a measurement would help us to determine whether peer discussions could have a positive influence on learning gain, and to understand better the

impact of why being more than three improves learning gain. This procedure is generally difficult to investigate in natural settings where several groups are working together in the same classroom.

Secondly, the specificity of the task used—a biology exercise—may also contribute to explaining the present results. Based on the group performance literature in social psychology, when a solution is proposed by a group member, and it is easily recognized as being correct (high demonstrability), the group outperforms the best performing individual (Laughlin and Ellis, 1986; Laughlin et al., 2002). In the present study, it is possible that if a student in a four-member group produced a correct drawing the first time round, his/her work could have been adopted by others the second time round, and consequently each group member improved his/her learning. Such a process is more likely to happen in ‘large’ groups than in dyads, or when group size increases (Laughlin et al., 1975). Although it is difficult to evaluate the demonstrability of the task, the biology exercise was designed by the teacher to be relatively complex for each student, and the demonstrability of the correct answer by the students was not obvious. On the other hand, if the exercise can be considered as a problem with low demonstrability, then the present results are not consistent with a study demonstrating that increasing group size contributes to decreasing performance for problems in which group members failed to recognize correct solutions proposed by their peers (Amir et al., 2018). An interpretation based on task demonstrability has to be considered with caution because most findings were obtained in social psychology studies measuring group performance in laboratory settings, and not individual learning. Rather than demonstrability, the complexity of the task should be considered as a crucial factor in engaging learners in meaningful discussions in order to perform an exercise correctly. For instance, it has been demonstrated that as task complexity increases learning as an individual becomes less effective and efficient than learning in a group of individuals (Kirschner et al., 2009). It is also possible to consider that, in four-member groups there is a greater likelihood of the presence of a group member with sufficient knowledge to catalyze learning improvement in members who are less knowledgeable than in smaller sized groups. Thus, future research should identify the level of knowledge of each group member before the task, their influence in the discussions and individual learning gain.

Beyond the effect of group size on learning gain, the present study revealed unexpected results about the moderating role of students’ gender. It appears that the benefits of peer discussions in “large” groups were greater for female than male participants. Indeed, results revealed that female participants in four-member groups significantly improve their learning more than those in the other conditions, and to a great extent than male participants involved in groups of the same size. For the male participants, no difference in learning gain was found between the experimental conditions, even though they tended to improve when they worked alone rather than when they discussed with their peers in groups of different sizes.

This moderating effect of gender could be interpreted in terms of the reduction of “social loafing” in groups. Indeed, a meta-analysis revealed that female participants are less likely to reduce their efforts for the group exhibiting a lower degree of ‘social loafing’ (Karau and Williams, 1993). On the contrary, by striving to be better than others through social comparisons (Buunk et al., 2007), male participants demonstrate a reverse pattern, regardless of the consequences on group performance (Karakowsky and Siegel, 1999). Of course, this interpretation is speculative because neither motivation, nor social comparisons within groups were measured, and no difference was found on learning gain among males related to group size.

Another interpretation may be proposed based on social role differences of male and female participants in cooperative or competitive contexts (Eagly, 1995; Datta Gupta et al., 2005; Niederle and Vesterlund, 2007; Michinov et al., 2009; Dohmen et al., 2011). According to the Social Role Theory (Eagly, 1987), males and females generally behave in ways that are consistent with their expected social roles, i.e., beliefs about gender differences in self-concepts. Based on this theory, women often have a more collectivist self-concept (cooperative, interdependent or communal), while men exhibit a more individualist self-concept (competitive, independent or agentic). A meta-analysis of gender difference in cooperation revealed that women were more cooperative than men in larger groups (Balliet et al., 2011).

Finally, our results were found in the context of technology-mediated learning, whereas prior research revealed mixed findings about the impact of group size on learning when students had to collaborate and discuss online with a large variety of technologies (Strijbos et al., 2004; Schellens and Valcke, 2006; Shaw, 2013; Sung et al., 2017; Sugai et al., 2019; Kim et al., 2020). On the contrary to previous studies, the technology-mediated learning environment used in the present study was based on tablets equipped with a stylus that allowed students to perform an exercise requiring a drawing to be produced a second time after having discussed their initial solutions with their peers in a face-to-face situation. In using such technology, collective feedback could be displayed to the whole class after the first drawing phase (see also Michinov et al., 2020), enabling all the work to be viewed in a “thumbnail” format. Although this format did not allow the students’ drawings to be seen in detail, an overview of the different drawings may have allowed them to support their thinking or discussions. However, because we could not record the students’ behaviors on video, and because there was insufficient time to administer a post-experimental questionnaire after the lesson, we ignore to what extent the students used the collective feedback to stimulate their discussion with their peers or their individual reflection.

CONCLUSION AND LIMITATIONS

The results of the present study revealed that discussing in four-member groups was more beneficial to learning improvement than discussing in smaller groups or no discussion with peers.

Thus, our findings tend to confirm that a group size of four individuals is the optimal configuration to improve peer learning. The present findings must be taken with caution because they are not without limitations like other quasi-experimental studies conducted in natural settings. One of the main limitations is the sample size yielding an interaction effect between gender and group size. Although some research is consistent with ours in claiming that the ceiling on group size should be four (Yetton and Bottger, 1983; Shimazoe and Aldrich, 2010), it would also be important in future studies to increase the group size to over four members. Another limitation concerns the lack of more in-depth investigations during or after the experiment, including collecting qualitative data from interviews or observations. Indeed, it is often difficult in field studies to take enough time with students and their teacher to observe or question them during a lesson, and particularly after the lesson. Although it would theoretically be possible to observe students performing the task and discussing with their peers during the experiment, such a method is difficult to apply practically and coding data problematic, in particular because many groups are working at the same time.

Far from being definitive, the present results need additional research to improve understanding of the impact of group size on learning, while also considering gender in group composition.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found below: <https://osf.io/5mehg/>.

REFERENCES

- Alexopoulou, E., and Driver, R. (1997). Small Group Discussions in Physics: Peer Interaction Modes in Pairs and Fours. *J. Res. Sci. Teach.* 33, 1099–1114. doi:10.1002/(SICI)1098-2736(199612)33:10<1099:AID-TEA4>3.0.CO;2-N
- Amir, O., Amir, D., Shahar, Y., Hart, Y., and Gal, K. (2018). The More the Merrier? Increasing Group Size May Be Detrimental to Decision-Making Performance in Nominal Groups. *PLOS ONE* 13, e0192213. doi:10.1371/journal.pone.0192213
- Apedoe, X. S., Ellefson, M. R., and Schunn, C. D. (2012). Learning Together while Designing: Does Group Size Make a Difference? *J. Sci. Educ. Technol.* 21, 83–94. doi:10.1007/s10956-011-9284-5
- Balliet, D., Li, N. P., Macfarlan, S. J., and Van Vugt, M. (2011). Sex Differences in Cooperation: A Meta-Analytic Review of Social Dilemmas. *Psychol. Bull.* 137, 881–909. doi:10.1037/a0025354
- Balta, N., Michinov, N., Balyimez, S., and Ayaz, M. F. (2017). A Meta-analysis of the Effect of Peer Instruction on Learning Gain: Identification of Informational and Cultural Moderators. *Int. J. Educ. Res.* 86, 66–77. doi:10.1016/j.ijer.2017.08.009
- Bray, R. M., Kerr, N. L., and Atkin, R. S. (1978). Effects of Group Size, Problem Difficulty, and Sex on Group Performance and Member Reactions. *J. Personal. Soc. Psychol.* 36, 1224–1240. doi:10.1037/0022-3514.36.11.1224
- Buunk, A. P., Cohen-Schotanus, J., and van Nek, R. H. (2007). Why and How People Engage in Social Comparison while Learning Social Skills in Groups. *Group Dyn. Theor. Res. Pract.* 11, 140–152. doi:10.1037/1089-2699.11.3.140

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

J-BC and NM worked in a collaborative fashion on this study. Both contributed to the study conception and design, and NM to the writing of the article.

FUNDING

This work was supported by the French Investment programme for the future (Digital innovation for educational excellence action). This research is a part of the ACTIF-eFRAN project (Digital training, research and animation area).

ACKNOWLEDGMENTS

We thank Coraline Bourges for her help in collecting data, and V ronique Payen for allowing us to conduct the study in her Life and Earth Sciences classes of the college Jacques Brel in Noyal-Sur-Vilaine (Brittany, France). Special thanks to Micka l Renault for his technical expertise in the programming of the KASSIS interactive learning environment, and Maxime Robert for his coordination of the ACTIF project.

- Caulfield, S. L., and Caroline, H. P. (2006). Teaching Social Science Reasoning and Quantitative Literacy: The Role of Collaborative Groups. *Teach. Sociol.* 34, 39–53. doi:10.1177/0092055X0603400104
- Clinton, V., and Kelly, A. E. (2020). Student Attitudes toward Group Discussions. *Active Learn. Higher Edu.* 21, 154–164. doi:10.1177/1469787417740277
- Craig, T. Y., and Kelly, J. R. (1999). Group Cohesiveness and Creative Performance. *Group Dyn. Theor. Res. Pract.* 3, 243–256. doi:10.1037/1089-2699.3.4.243
- Datta Gupta, N., Poulsen, A., and Villeval, M.-C. (2005). Male and Female Competitive Behavior - Experimental Evidence. *GATE Working Paper No. W.P.05-12* doi:10.2139/ssrn.906766
- Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *J. Eur. Econ. Assoc.* 9, 522–550. doi:10.1111/j.1542-4774.2011.01015.x
- Eagly, A. H., Karau, S. J., and Makhijani, M. G. (1995). Gender and the Effectiveness of Leaders: A Meta-Analysis. *Psychol. Bull.* 117, 125–145. doi:10.1037/0033-2909.117.1.125
- Eagly, A. H. (1987). *Sex Differences in Social Behavior: A Social-Role Interpretation*. Hillsdale, NJ: Erlbaum.
- Egerbladh, T. (1976). The Function of Group Size and Ability Level on Solving a Multidimensional Complementary Task. *J. Personal. Soc. Psychol.* 34, 805–808. doi:10.1037/0022-3514.34.5.805
- Hackman, J. R., Jones, L. E., and McGrath, J. E. (1967). A Set of Dimensions for Describing the General Properties of Group-Generated Written Passages. *Psychol. Bull.* 67, 379–390. doi:10.1037/h0024647
- Hamburger, H., Guyer, M., and Fox, J. (1975). Group Size and Cooperation. *J. Conflict Resolut.* 19, 503–531. doi:10.1177/002200277501900307

- Hill, G. W. (1982). Group versus Individual Performance: Are N+1 Heads Better Than One? *Psychol. Bull.* 91, 517–539. doi:10.1037/0033-2909.91.3.517
- Ingham, A. G., Levinger, G., Graves, J., and Peckham, V. (1974). The Ringelmann Effect: Studies of Group Size and Group Performance. *J. Exp. Soc. Psychol.* 10, 371–384. doi:10.1016/0022-1031(74)90033-x
- Kagan, S. (1992). *Cooperative Learning*. 2nd ed. San Clemente, CA: Kagan Cooperative Learning.
- Karakowsky, L., and Siegel, J. P. (1999). The Effects of Proportional Representation and Gender Orientation of the Task on Emergent Leadership Behavior in Mixed-Gender Work Groups. *J. Appl. Psychol.* 84, 620–631. doi:10.1037/0021-9010.84.4.620
- Karau, S. J., and Williams, K. D. (1993). Social Loafing: A Meta-Analytic Review and Theoretical Integration. *J. Personal. Soc. Psychol.* 65, 681–706. doi:10.1037/0022-3514.65.4.681
- Kim, N. J., Belland, B. R., Lefler, M., Andreasen, L., Walker, A., and Axelrod, D. (2020). Computer-based Scaffolding Targeting Individual versus Groups in Problem-Centered Instruction for Stem Education: Meta-Analysis. *Educ. Psychol. Rev.* 32, 415–461. doi:10.1007/s10648-019-09502-3
- Kirschner, F., Paas, F., and Kirschner, P. A. (2009). A Cognitive Load Approach to Collaborative Learning: United Brains for Complex Tasks. *Educ. Psychol. Rev.* 21, 31–42. doi:10.1007/s10648-008-9095-2
- Knight, J. K., and Brame, C. J. (2018). Peer Instruction. *CBE Life Sci. Educ.* 17, fe5. doi:10.1187/cbe.18-02-0025
- Komorita, S. S., and Lapworth, C. W. (1982). Cooperative Choice Among Individuals versus Groups in an N-Person Dilemma Situation. *J. Personal. Soc. Psychol.* 42, 487–496. doi:10.1037/0022-3514.42.3.487
- Latan , B., Williams, K., and Harkins, S. (1979). Many Hands Make Light the Work: The Causes and Consequences of Social Loafing. *J. Personal. Soc. Psychol.* 37, 822–832. doi:10.1037/0022-3514.37.6.822
- Laughlin, P. R., Hatch, E. C., Silver, J. S., and Boh, L. (2006). Groups Perform Better Than the Best Individuals on Letters-To-Numbers Problems: Effects of Group Size. *J. Pers. Soc. Psychol.* 90, 644–651. doi:10.1037/0022-3514.90.4.644
- Laughlin, P. R., Bonner, B. L., and Miner, A. G. (2002). Groups Perform Better Than the Best Individuals on Letters-To-Numbers Problems. *Organ. Behav. Hum. Decis. Process.* 88, 605–620. doi:10.1016/S0749-5978(02)00003-1
- Laughlin, P. R., and Ellis, A. L. (1986). Demonstrability and Social Combination Processes on Mathematical Intellectual Tasks. *J. Exp. Soc. Psychol.* 22, 177–189. doi:10.1016/0022-1031(86)90022-3
- Laughlin, P. R., Kerr, N. L., Davis, J. H., Halff, H. M., and Marciniak, K. A. (1975). Group Size, Member Ability, and Social Decision Schemes on an Intellectual Task. *J. Personal. Soc. Psychol.* 31, 522–535. doi:10.1037/h0076474
- Levine, J. M., and Moreland, R. L. (2004). Collaboration: The Social Context of Theory Development. *Pers Soc. Psychol. Rev.* 8, 164–172. doi:10.1207/s15327957pspr0802_10
- Littlepage, G. E. (1991). Effects of Group Size and Task Characteristics on Group Performance: A Test of Steiner's Model. *Pers. Soc. Psychol. Bull.* 17, 449–456. doi:10.1177/0146167291174014
- Littlepage, G. E., and Silbiger, H. (1992). Recognition of Expertise in Decision-Making Groups. *Small Group Res.* 23, 344–355. doi:10.1177/1046496492233005
- Lohman, M. C., and Finkelstein, M. (2000). Designing Groups in Problem-Based Learning to Promote Problem-Solving Skills and Self-Directedness. *Instr. Sci.* 28, 291–307. doi:10.1023/A:1003927228005
- Lorge, I., and Solomon, H. (1959). Individual Performance and Group Performance in Problem Solving Related to Group Size and Previous Exposure to the Problem. *J. Psychol.* 48, 107–114. doi:10.1080/00223980.1959.9916346
- Lou, Y., Abrami, P. C., and d'Apollonia, S. (2001). Small Group and Individual Learning with Technology: A Meta-Analysis. *Rev. Educ. Res.* 71, 449–521. doi:10.3102/00346543071003449
- Mazur, E. (1997). *Peer Instruction: A User's Manual Series in Educational Innovation*. Upper Saddle River, NJ: Prentice-Hall.
- McDonald, B. A., Larson, C. O., Dansereau, D. F., and Spurlin, J. E. (1985). Cooperative Dyads: Impact on Text Learning and Transfer. *Contemp. Educ. Psychol.* 10, 369–377. doi:10.1016/0361-476X(85)90033-5
- Melero, J., Hern andez-Leo, D., and Manatunga, K. (2015). Group-based mobile Learning: Do Group Size and Sharing mobile Devices Matter? *Comput. Hum. Behav.* 44, 377–385. doi:10.1016/j.chb.2014.11.078
- Michinov, E., Michinov, N., and Huguet, P. (2009). Effects of Gender Role and Task Content on Performance in Same-Gender Dyads: Transactive Memory as a Potential Mediator. *Eur. J. Psychol. Educ.* 24, 155–168. doi:10.1007/BF03173008
- Michinov, N., Anquetil,  ., and Michinov, E. (2020). Guiding the Use of Collective Feedback Displayed on Heatmaps to Reduce Group Conformity and Improve Learning in Peer Instruction. *J. Comput. Assist. Learn.* 36, 1026–1037. doi:10.1111/jcal.12457
- Michinov, N., Michinov, E., and Toczec-Capelle, M.-C. (2004). Social Identity, Group Processes, and Performance in Synchronous Computer-Mediated Communication. *Group Dyn. Theor. Res. Pract.* 8, 27–39. doi:10.1037/1089-2699.8.1.27
- Michinov, N., Morice, J., and Ferri eres, V. (2015). A Step Further in Peer Instruction: Using the Stepladder Technique to Improve Learning. *Comput. Edu.* 91, 1–13. doi:10.1016/j.compedu.2015.09.007
- Moreland, R. L. (2010). Are Dyads Really Groups? *Small Group Res.* 41, 251–267. doi:10.1177/1046496409358618
- Morice, J., Michinov, N., Delaval, M., Sideridou, A., and Ferri eres, V. (2015). Comparing the Effectiveness of Peer Instruction to Individual Learning during a Chromatography Course. *J. Comp. Assist. Learn.* 31, 722–733. doi:10.1111/jcal.12116
- Needham, R. (1987). *Teaching Strategies for Developing Understanding in Science*. Leeds, UK: Centre for Studies in Science and Mathematics Education, Children's Learning in Science Project.
- Niederle, M., and Vesterlund, L. (2007). Do women Shy Away from Competition? Do Men Compete Too Much? *Q. J. Econ.* 122, 1067–1101. doi:10.1162/qjec.122.3.1067
- Peltokorpi, J., and Niemi, E. (2019). Effects of Group Size and Learning on Manual Assembly Performance: An Experimental Study. *Int. J. Prod. Res.* 57, 452–469. doi:10.1080/00207543.2018.1444810
- Petty, R. E., Harkins, S. G., Williams, K. D., and Latane, B. (1977). The Effects of Group Size on Cognitive Effort and Evaluation. *Pers. Soc. Psychol. Bull.* 3, 579–582. doi:10.1177/014616727700300406
- Prince, M. (2004). Does Active Learning Work? A Review of the Research. *J. Eng. Educ.* 93, 223–231. doi:10.1002/j.2168-9830.2004.tb00809.x
- Relling, A. E., and Giuliodori, M. J. (2015). Effect of Peer Instruction on the Likelihood for Choosing the Correct Response to a Physiology Question. *Adv. Physiol. Educ.* 39, 167–171. doi:10.1152/advan.00092.2014
- Richey, J. E., Nokes-Malach, T. J., and Cohen, K. (2018). Collaboration Facilitates Abstract Category Learning. *Mem. Cognit.* 46, 685–698. doi:10.3758/s13421-018-0795-7
- Schell, J. A., and Butler, A. C. (2018). Insights from the Science of Learning Can Inform Evidence-Based Implementation of Peer Instruction. *Front. Educ.* 3, 33. doi:10.3389/educ.2018.00033
- Schellens, T., and Valcke, M. (2006). Fostering Knowledge Construction in university Students through Asynchronous Discussion Groups. *Comput. Edu.* 46, 349–370. doi:10.1016/j.compedu.2004.07.010
- Schultz, B. G. (1989). *Communicating in the Small Group: Theory and Practice*. New York: Harper & Row.
- Schultze, T., Mojzisch, A., and Schulz-Hardt, S. (2012). Why Groups Perform Better Than Individuals at Quantitative Judgment Tasks: Group-To-Individual Transfer as an Alternative to Differential Weighting. *Organ. Behav. Hum. Decis. Process.* 118, 24–36. doi:10.1016/j.obhdp.2011.12.006
- Seta, J. J., Paulus, P. B., and Schkade, J. K. (1976). Effects of Group Size and Proximity under Cooperative and Competitive Conditions. *J. Personal. Soc. Psychol.* 34, 47–53. doi:10.1037/0022-3514.34.1.47
- Shaw, M. E. (1963). Scaling Group Tasks: A Method for Dimensional Analysis. Technical report No. 1. Gainesville, Florida: Department of Psychology, University of Florida, Office of Naval Research Contract NR, 170–266. Nonr-580(11).
- Shaw, R.-S. (2013). The Relationships Among Group Size, Participation, and Performance of Programming Language Learning Supported with Online Forums. *Comput. Edu.* 62, 196–207. doi:10.1016/j.compedu.2012.11.001
- Shimazoe, J., and Aldrich, H. (2010). Group Work Can Be Gratifying: Understanding & Overcoming Resistance to Cooperative Learning. *Coll. Teach.* 58, 52–57. doi:10.1080/87567550903418594
- Slavin, R. E. (1987). Ability Grouping and Student Achievement in Elementary Schools: A Best-Evidence Synthesis. *Rev. Educ. Res.* 57, 293–336. doi:10.3102/00346543057003293

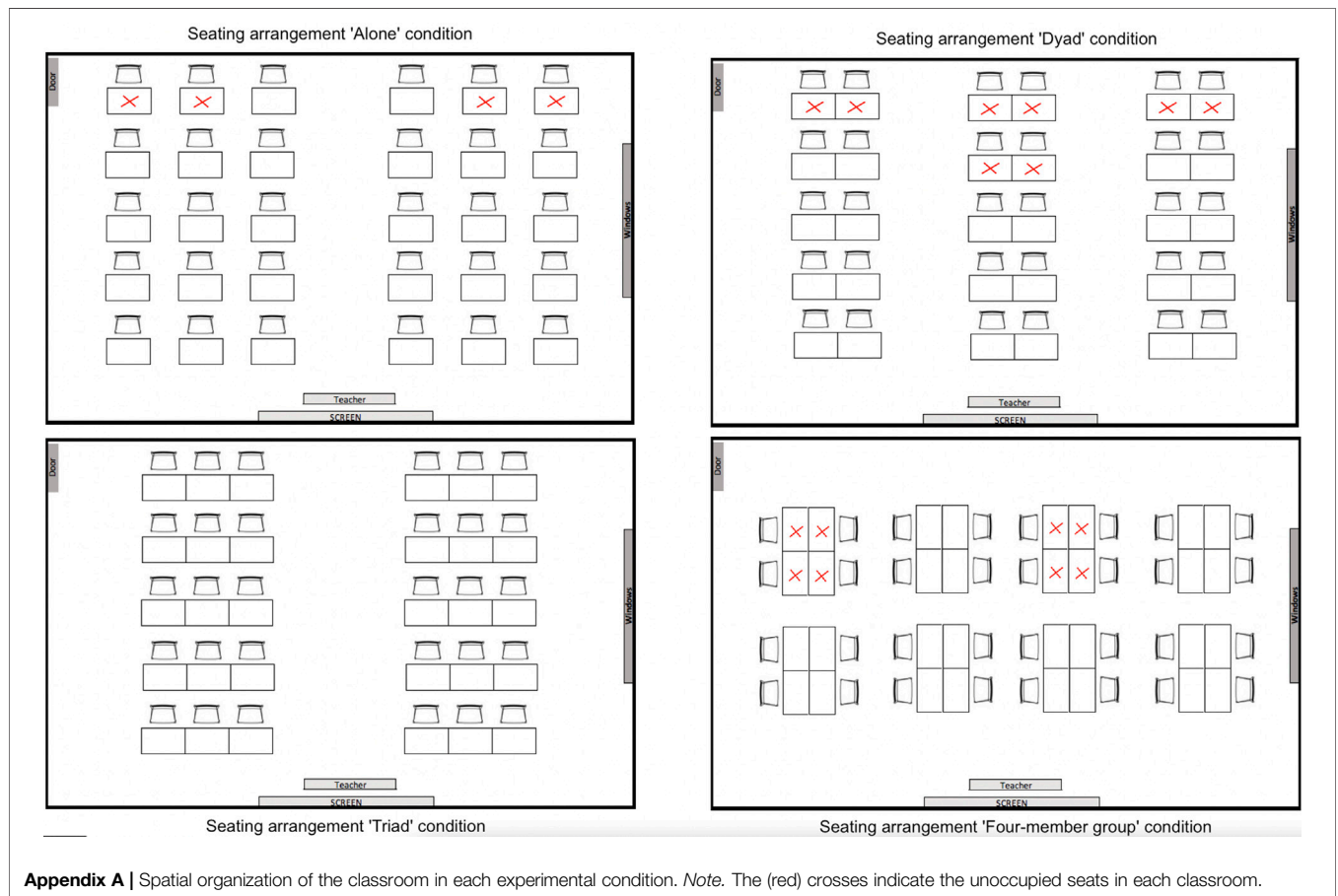
- Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., et al. (2009). Why Peer Discussion Improves Student Performance on In-Class Concept Questions. *Science* 323, 122–124. doi:10.1126/science.1165919
- Steiner, I. D. (1972). *Group Process and Productivity*. New York: Academic Press.
- Strijbos, J.-W., Martens, R. L., Jochems, W. M. G., and Broers, N. J. (2004). The Effect of Functional Roles on Group Efficiency. *Small Group Res.* 35, 195–229. doi:10.1177/1046496403260843
- Sugai, M., Horita, T., and Wada, Y. (2019). Optimal Group Size for High School Students' Collaborative Argumentation Using SNS for Educational Purposes. *Int. J. Learn. Technol. Learn. Environ.* 2, 35–53. doi:10.52731/ijltle.v2.i2.410
- Sung, Y. T., Yang, J. M., and Lee, H. Y. (2017). The Effects of mobile-computer-supported Collaborative Learning: Meta-Analysis and Critical Synthesis. *Rev. Educ. Res.* 87, 768–805. doi:10.3102/0034654317704307
- Tasca, G. A. (2020). What Is Group Dynamics? *Group Dyn. Theor. Res. Pract.* 24, 1–5. doi:10.1037/gdn0000115
- Taylor, D. W., and Faust, W. L. (1952). Twenty Questions: Efficiency in Problem Solving as a Function of Size of Group. *J. Exp. Psychol.* 44, 360–368. doi:10.1037/h0054376
- Thomas, E. J., and Fink, C. F. (1963). Effects of Group Size. *Psychol. Bull.* 60, 371–384. doi:10.1037/h0047169
- Topping, K., Buchs, C., Duran, D., and van Keer, H. (2017). *Effective Peer Learning: From Principles to Practical Implementation*. 1st ed. New York: Routledge. doi:10.4324/9781315695471
- Vickrey, T., Rosploch, K., Rahmanian, R., Pilarz, M., and Stains, M. (2015). Research-based Implementation of Peer Instruction: A Literature Review. *CBE Life Sci. Educ.* 14, es3. doi:10.1187/cbe.14-11-0198
- Webb, N. M. (1989). Peer Interaction and Learning in Small Groups. *Int. J. Educ. Res.* 13, 21–39. doi:10.1016/0883-0355(89)90014-1
- Webb, N. M. (1982). Student Interaction and Learning in Small Groups. *Rev. Educ. Res.* 52, 421–445. doi:10.3102/00346543052003421
- Wheelan, S. A. (2009). Group Size, Group Development, and Group Productivity. *Small Group Res.* 40, 247–262. doi:10.1177/1046496408328703
- Wiley, J., and Jensen, M. (2006). "When Three Heads Are Better Than Two," in *Proceedings, CogSci 2006* (Vancouver, CA: Cognitive Science Society).
- Williams, K. D. (2010). Dyads Can Be Groups (And Often Are). *Small Group Res.* 41, 268–274. doi:10.1177/1046496409358619
- Yetton, P., and Bottger, P. (1983). The Relationships Among Group Size, Member Ability, Social Decision Schemes, and Performance. *Organ. Behav. Hum. Perform.* 32, 145–159. doi:10.1016/0030-5073(83)90144-7

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright   2021 Corr ge and Michinov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX



RÉPONSES

Appendix B | Overview of the collective feedback displayed on the screen in a “thumbnail” format.