Computers & Education 91 (2015) 1–13

Contents lists available at ScienceDirect

Computers & Education

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

A step further in Peer Instruction: Using the Stepladder technique to improve learning

Nicolas Michinov [a,*], Julien Morice [b], Vincent Ferrières [c]

[a] Université Rennes 2 (CRPCC, EA 1285), France
[b] Université européenne de Bretagne, France
[c] Ecole Nationale Supérieure de Chimie de Rennes (CNRS, UMR 6226), France

ARTICLE INFO

Article history:
Received 29 May 2015
Received in revised form 10 September 2015
Accepted 11 September 2015
Available online 14 September 2015

Keywords:
Peer Instruction
Stepladder technique
Learning gains
Instructional method
Higher education

ABSTRACT

Peer Instruction (PI) is an instructional strategy for engaging students during class through a structured questioning process that improves the learning of the concepts of fundamental sciences. Although all students are supposedly engaged in discussions with their peers during Peer Instruction, the learning gains generally remain at a medium level, suggesting a lack of participation of certain students who do not benefit from social interactions. The present study examined whether the Stepladder technique might optimize the Peer Instruction method and increase learning gains. With this technique, students enter a group sequentially, forcing every group member to participate in discussions. Eighty-four chemistry students were asked to answer easy and difficult multiple-choice questions before and after being randomly assigned to one of three instructional conditions during a chromatography lesson (Classic PI vs. Stepladder PI vs. Individual Instruction without any discussion with peers). As predicted, results showed that learning gains were greatest in the Stepladder PI group, and that this effect was mainly observed for difficult questions. Results also revealed higher perceived satisfaction when students had to discuss the questions with their peers than when they were not given this possibility. By extending the Stepladder technique to higher education, these findings offer a step forward in the Peer Instruction literature, showing how it can enhance learning gains.

1. Introduction

Peer Instruction (PI) is an interactive student-centred instructional strategy for engaging students in class through a structured questioning process that improves the learning of the concepts of fundamental sciences. Although a number of studies have revealed that the Peer Instruction method produces greater learning gains than more traditional instructional methods such as lectures, the gains generally remain at a medium level (e.g., Crouch, Watkins, Fagen, & Mazur, 2007; Crouch & Mazur, 2001; Hake, 1998; Mayer et al., 2009; Mazur, 1997). One reason may be found in the premise that all students in groups are actively engaged in fruitful discussions with their peers during a PI session. However, evidence suggests that a substantial number of individuals are reluctant to communicate their ideas verbally in group and classroom settings for fear of being judged by others (e.g., McCroskey & Beatty, 1984; Micari & Pazos, 2014). Some experimentally tested techniques, such

* Corresponding author. Université Rennes 2 (CRPCC, EA 1285), Place du Recteur Henri Le Moal, CS 24307, 35043 Rennes Cedex, France.
E-mail addresses: nicolas.michinov@univ-rennes2.fr (N. Michinov), julienmorice@gmail.com (J. Morice), vincent.ferrieres@ensc-rennes.fr (V. Ferrières).

http://dx.doi.org/10.1016/j.compedu.2015.09.007
0360-1315/

as the Stepladder technique (e.g., Rogelberg, Barnes-Farrell, & Lowe, 1992; Rogelberg & O'Connor, 1998), have been shown to increase participation in groups because the entry of members is structured in such a way that every individual is required to contribute to the discussion in turn. To our knowledge, no studies have examined the use of this technique outside the laboratory, particularly in the sciences in higher education where students have to learn difficult concepts. As this technique is known to boost participation and performance in groups, it is reasonable to consider that it may also optimize the Peer Instruction method. Consequently, the purpose of the present study was to examine whether the Stepladder technique, known to improve performance in groups by actively engaging all members, might optimize the efficacy of Peer Instruction.

1.1. The Peer Instruction method to improve learning

1.1.1. Main findings in Peer Instruction and related socio-constructivist methods

The Peer Instruction method was initially developed by the Harvard physicist Eric Mazur (1997) for a course entitled “Physics for Non-Majors”. This method changes the traditional lecture format to one in which the instructor poses questions during the lecture, in order to engage students actively in discussions with their peers and focus their attention on central concepts (e.g., Crouch et al., 2007; Crouch & Mazur, 2001; Mazur, 1997). By providing opportunities for students to address difficult aspects of course material, Peer Instruction enables them to learn from each other by sharing ideas and solutions. It was initially introduced to promote interactions between physics students, who had to answer multiple-choice questions of the Force Concept Inventory (FCI; Hestenes, Wells, & Swackhamer, 1992) using personal response systems or clickers.[1] For more than a decade, Peer Instruction has been recognized by scientists as an efficient instructional method that enhances deeper learning by encouraging students to interact with their peers instead of staying passive (e.g., Crouch et al., 2007; Crouch & Mazur, 2001; Mazur, 1997). Although different versions of this method have been used, it is usually organized as follows: students vote individually on the correct answers to multiple-choice questions using clickers; if there is a moderate percentage of correct answers, the teacher asks the students to discuss the questions with their neighbours for 2–3 min; they then vote individually again on the same questions. The sequence generally ends with a whole-class discussion in which the instructor provides explanations about the concepts. Although this is the most generally recommended way of using the Peer Instruction method (Mazur, 1997), a wide variety of variants has been observed, and a study revealed that less than 12.8% of the teachers using Peer Instruction do so as it was originally designed to be implemented (Dancy & Henderson, 2010). More broadly, a typical Peer Instruction sequence resembles a 'Think-Pair-Share' learning strategy (e.g., Watkins & Mazur, 2010), whereby students have to think individually about questions before sharing ideas and solutions with their classmates, either in pairs or in three- or four-member groups.

Peer Instruction is thus similar to other learning methods based on interactions with peers, such as Peer feedback (e.g., Van Zundert, Sluijsmans, & Van Merriënboer, 2010) and Peer-led Team Learning (e.g., Hockings, DeAngelis, & Frey, 2008). The former has been used essentially to improve students' writing skills, and involves equal-status learners giving comments/feedback to their peers. This method has been shown to improve several aspects of learning (e.g., Gielen, Peeters, Dochy, Onghena, & Struyven, 2010). Similarly, the Peer-led Team Learning method has been used to improve learning in chemistry classes, and involves good students becoming leaders of small groups of students, helping them solve problems through discussion and debate. Studies have demonstrated learning improvements with these methods compared to traditional lectures (e.g., Hockings et al., 2008; Lewis & Lewis, 2005; Wamser, 2006). The benefits of peer discussions for learning have also been demonstrated in a large variety of other instructional methods, such as reciprocal teaching (e.g., Rosenshine & Meister, 1994), peer tutoring (e.g., Fantuzzo, Riggio, Connelly, & Dimeff, 1989; Palincsar & Brown, 1984), and (Computer-Supported) collaborative learning (e.g., Goodyear, Jones, & Thompson, 2014), including 'collaboration scripts' (e.g., Kobbe et al., 2007; Weinberger, Fischer, & Mandl, 2002). Many other instructional methods based on cooperative learning have been described in the literature (see Johnson & Johnson, 2008; Slavin, 2011), including Student Teams-Achievement Divisions (STAD), Teams–Games–Tournaments (TGT), Group Investigation (GI), and Jigsaw Grouping. Like Peer Instruction, all these methods and instructional strategies are based on a social constructivist approach to learning, in which social interaction plays a crucial role in the construction of knowledge, and where discussion and collaboration between peers have a positive impact on learning (e.g., Dillenbourg, Baker, Blaye, & O'Malley, 1996; Doise & Mugny, 1984; Springer, Stanne, & Donovan, 1999). All these methods indicate that learning is not only the product of individual cognitive processes, but that it can be improved through group discussion, helping students develop a deeper understanding of what they are learning.

Since it was first developed by Mazur (1997), the Peer Instruction method has been widely recognized by scientists to have a positive impact on learning, particularly for difficult concepts of fundamental sciences such as physics and chemistry (e.g., Crouch & Mazur, 2001; Crouch et al., 2007; Fagen, Crouch, & Mazur, 2002; Lasry, Mazur, & Watkins, 2008), although it is not limited to these two disciplines (e.g., Rao & DiCarlo, 2000; Zingaro & Porter, 2014). It has generally been demonstrated to be more effective than traditional methods without peer interaction, such as lectures attended by a large number of students (e.g., Hake, 1998; Mayer et al., 2009) and individualized learning (e.g., Jones, Antonenko, & Greenwood, 2012; Lasry, Charles, Whittaker, & Lautman, 2009). In contrast to traditional learning methods, Peer Instruction was designed to involve all the students in solving problems through peer discussion, thereby improving the learning of difficult concepts (Zingaro & Porter,

[1] Personal response systems, also named “clickers”, allow students to answer questions in class providing real-time and anonymous feedback that can be used by students to share their ideas and solutions with others (e.g., Banks, 2006; Lantz, 2010).

2014). A number of researchers have demonstrated that the proportion of correct answers given by students increases significantly after a period of discussion with their peers (e.g., Hake, 1998; Singh, 2005; Smith, Wood, Adams, Wieman, Knight et al., 2009), even if all the initial responses of the group were wrong (e.g., Porter, Bailey-Lee, Simon, & Zingaro, 2011; Singh, 2005; Smith et al., 2009). In a meta-analysis that aimed to measure the learning gains of 6542 students enrolled in introductory physics courses using interactive or traditional methods over a semester, Hake (1998) gathered data of pre- and post-test multiple-choice questions assessing students' understanding of the basic concepts of Newtonian physics. Results demonstrated that 85% of the interactive courses and none of the traditionally taught courses yielded a medium learning gain (g) ranging from 0.3 to 0.7. More recently, a survey was carried out among college and university physics instructors who provided pre- and post-test data to assess the learning gains of students (Fagen et al., 2002). The results showed that the Peer Instruction courses had a class average gain of 0.39, and that a large majority of these courses came within the medium gain range. Similar results were recently found in engineering and mathematics as well as physics (Freeman, Eddy, McDonough, Smith, Okoroafor et al., 2014).

1.1.2. Some limitations of Peer Instruction

The learning gain from Peer Instruction is supposedly the result of discussions between peers who, by comparing and explaining their views, produce a better response (e.g., Weber, Maher, Powell, & Lee, 2008). It is widely accepted that Peer Instruction involves a structured questioning process that engages every student in the class (e.g., Crouch et al., 2007). However, there is no clear empirical evidence supporting the view that every student is active and participates in the discussion. Although a large number of studies have demonstrated that the Peer Instruction method improves learning in comparison to more traditional instructional methods, the learning gains remain at a medium level, and a difficulty to exceed this level has frequently been observed (e.g., Crouch & Mazur, 2001; Hake, 1998). This finding of a relatively limited scope of the Peer Instruction method may be partly explained by a lack of participation of some students; this is supported by empirical data and observations of Peer Instruction sessions. First, some researchers have observed that only a few students participate in peer discussions, while others remain silent or contribute less actively (e.g., Brooks & Koretsky, 2011; Lucas, 2009). Secondly, in a study measuring 'conversation bias', i.e. the difference between the proportion of statements made by one or other partner during a discussion, James (2006) found that in classes where little credit is given for incorrect responses, as is generally the case in most educational settings, students with greater knowledge dominate the discussions, while those with less knowledge remain more passive. It is reasonable to assume that this bias may also appear during a Peer Instruction session with larger groups. Finally, as highlighted in a survey of physics instructors using the Peer Instruction method, a substantial number admitted that the "challenge is the difficulty in fully engaging students in class discussions" (Fagen et al., 2002, p. 208). In that study, conducted by the advocates of Peer Instruction themselves, nearly half the instructors who mentioned this challenge suggested ways of overcoming this difficulty, such as moving round the classroom during the peer discussion, or guiding students and encouraging them to share their ideas and solutions with each other. To date, these possibilities have not been explored, and any instructional procedures that might engage all, and not just some, participants in discussions need to be examined. One possibility is to use a guided and structured procedure to organize discussions during a Peer Instruction session. As suggested by Lucas (2009), a formalized written procedure could be used, in which students are first asked to write down their own arguments before any discussion has taken place, and then read out what they have written. Using this type of procedure was found to improve students' participation in the Peer Instruction session and their comprehension. More importantly, by forcing each group member to present his or her own reasoning, this procedure improved the correct response rate and reduced the number of erroneous arguments. Unfortunately, this type of written procedure has not been experimentally compared with the classic Peer Instruction procedure. As it is very similar to another method, the Stepladder technique, in which members of the group are asked in turn to present their arguments verbally, we can reasonably suppose that the latter could also enhance the effectiveness of learning compared to the standard Peer Instruction procedure.

1.2. Extending the Stepladder technique to Peer Instruction to improve learning

A problem-solving structure to overcome the problem of unequal participation in groups, named the Stepladder technique, was proposed by Rogelberg and collaborators (e.g., Rogelberg et al., 1992; Rogelberg & O'Connor, 1998).

The Stepladder technique begins with a discussion between two randomly chosen group members (named the initial core group) working together on a problem and sharing their views and ideas. Other group members are added to the discussion one at a time and present their views on the problem before hearing the core group's preliminary solution. The group members take the time to listen to, and to understand, the propositions of each new member. The process continues until all the group members' views have been presented and discussed, and ends when discrepancies between the different views have been solved. With this technique, group members must have enough time to consider the problem or decision on their own, present their views to the group, and discuss all views with the group at each step. The main advantage of the Stepladder technique is that it prevents deleterious group dynamics in several ways, for example by facilitating communication of all group members instead of only a few, giving each group member the opportunity to present and share their knowledge with others, reducing social loafing because members cannot hide behind others' contributions, fostering critical thinking and controversy in groups, generating better performance, etc.

To date, this technique has only been tested in laboratory settings using simulation tasks where individuals in four-member groups have to rank objects by order of importance for their survival. Results generally show a relative success of this technique (Orpen, 1997; Rogelberg et al., 1992; Rogelberg & O'Connor, 1998), including when using synchronous communication technologies such as audio (Rogelberg, O'Connor, & Sederburg, 2002) and textual conferencing systems (Thompson & Coovert, 2002). The Stepladder technique has also been found to have a positive effect on satisfaction (Rogelberg et al., 1992). As far as we know, the only researchers who have failed to replicate the positive effects generally observed in the laboratory are Winquist and Franz (2008) who highlighted some limitations of the technique such as experimenter, sample, and task differences. They also suggested that future research should “examine whether interventions such as the Stepladder Technique work better when using richer and more complex tasks than when using more simplistic ones” (Winquist & Franz, 2008, p. 266). This echoes the point raised by Rogelberg and collaborators that “a common critique of most lab-based group research is whether it will generalize to the world of work where real employees are doing real tasks” (Rogelberg et al., 2002, p. 999). Beyond the world of work, and as far as we know, the use of this technique in learning environments has never been examined, particularly in relation to Peer Instruction in which all the students are supposed to join in a group discussion to answer questions of varying difficulty. The aim of the present study was thus to extend this technique to learning environments in order to optimize the Peer Instruction method, and consequently, to improve learning gains.

1.3. Peer Instruction or individual Instruction to improve learning gains?

To be sure that learning gains are due to interaction between peers, the Peer Instruction method, either in its classic form or structured around the Stepladder technique, must be compared to a condition where students have to work individually without interacting with their peers. To date, the few studies that have examined the effectiveness of the Peer Instruction method compared to that of individuals working alone have produced mixed results. Some studies found that group discussion gives better results than individual instruction methods (e.g., Jones et al., 2012; Smith et al., 2009; Zingaro & Porter, 2014). For example, Jones and collaborators (2012) compared the Peer Instruction method to individualized student response system-based instruction in a large undergraduate science course in order to examine their respective impacts on learner motivation, regulation of cognition, and learning transfer. In that study, participants in the Peer Instruction group first responded to conceptual questions, discussed their responses in small groups, and then provided a revised response to the question. Participants in the other group gave individualized responses to the same questions. Amongst other things, the results revealed that learning transfer was better in the Peer Instruction group than in the Individual Instruction group. However, the learning gain benefits of engaging in peer discussions were not tested. In a similar study comparing Peer Instruction to individual learning in a blended learning environment where students have access to online resources, no difference in learning gain was observed between the two conditions (Morice, Michinov, Delaval, Sideridou, & Ferrières, in press). In another study, Lasry et al. (2009) showed that learning gains were greater when students were asked to discuss their choice with a peer than in other conditions of individual instruction (reflect for a minute without discussion, perform a distraction task without any discussion or reflection). Finally, other studies also revealed that peer-to-peer interactions are more beneficial than interacting with technology alone, and collaboration in groups leads to better learning outcomes compared to individual learning (e.g., Johnson & Johnson, 2008; Lou, Abrami, & d'Appolonia, 2001; Perlmutter, Behrend, Kuo, & Muller, 1989).

1.4. Overview of the present study and hypotheses

In the present study, chemistry students were asked to answer 10 multiple-choice questions of varying difficulty (easy and difficult) based on a pretest, using clickers twice during a 90-min chromatography lesson in an active learning classroom (Baepler, Walker, & Driessen, 2014). They were randomly assigned to one of three instructional groups: Classic Peer Instruction, Stepladder Peer Instruction, or Individual Instruction without any discussion with peers. Learning gains were calculated from their answers at the beginning and at the end of the lesson. Two hypotheses were put forward, based on previous findings in the Peer Instruction and Stepladder technique literature.

Hypothesis 1. As the Stepladder technique has been shown to be more efficient and satisfactory than the conventional unstructured technique in which group members join the discussion simultaneously (Orpen, 1997; Rogelberg et al., 1992; Thompson & Coovert, 2002), it was expected that structuring the technique so that every group member was forced to participate in turn (Stepladder Peer Instruction condition) would enable students to learn more than in the other conditions (Classic Peer Instruction and Individual Instruction conditions).

Hypothesis 2. As the Peer Instruction method has been shown to have a positive impact on learning in the sciences (Smith et al., 2009), particularly for difficult questions (Zingaro & Porter, 2014), it was expected that learning gains would be greater in the Stepladder Peer Instruction condition than in the other conditions for difficult questions. No difference between groups was expected for easy questions.[2]

[2] As the literature provides mixed results regarding the superiority of the Peer Instruction method over individualized learning where students are actively involved in their learning (e.g., Jones et al., 2012; Lasry et al., 2009; Morice et al., in press), it was difficult to hypothesize that students in Classic Peer Instruction would have greater learning gains than those in Individual Instruction.

2. Method

2.1. Participants

Participants were 84 students (38 women and 46 men) aged 20–23 years ($M = 21.2$, $SD = 0.71$) in a chemical engineering school. None of them had used a clicker-based system in their studies, and they had never experienced a Peer Instruction method.

2.2. Materials

2.2.1. Clicker device

The experimental equipment consisted of a set of clickers and a receiver connected to a computer, linked in turn to a video projector. A single software package (CPS Pulse, e-instruction® by Turning technologies) enabled us to create and administer interactive PowerPoint® multiple-choice questions and to view participation rates and results in real time. More importantly, it provided the students with feedback on their responses in the form of a histogram displayed on the screen at the front of the classroom. The data could be exported in the form of a spreadsheet for further statistical analysis.

2.2.2. Multiple-choice questions

A set of conceptual chromatography questions was developed for a 90-min lesson. The questions were extracted from a database to be evaluate in a pilot study among a similar sample of students. Based on the pilot study, a number of questions were excluded because of ceiling and floor effects, and ten multiple-choice questions with varying levels of difficulty were retained (see Appendix for a sample of questions). As several studies have shown that questions that are too easy limit the learning potential of Peer Instruction (Crouch et al., 2007; Zingaro & Porter, 2014), questions with a correct response rate of 80% or above in the first test in the pilot study were dropped. It was also decided to drop the questions that were "too difficult", namely those that were correctly answered by fewer than 20% of students. As shown in Table 1, the percentages of correct answers in the first test for the ten multiple-choice questions of the pilot study are almost identical to those observed in the present study, suggesting that the choice of questions based on a difficulty criterion was relevant. Indeed, the results of the pilot study revealed 64.8% of correct answers for the five easy questions retained, compared to 28.8% for the five difficult questions (in the present study, 56.6% and 26.4%, respectively).

2.3. Procedure

The study took place during the first semester in a large classroom arranged to promote active learning. In many active learning classrooms (e.g., Baepler et al., 2014), the tables are arranged in clusters to support discussion and work in small groups, generally in conjunction with additional learning technologies and Internet access to retrieve resources. In the present study, four students were seated at each table with a PC, and one side of the table was kept free so that no one had their back to the teacher's desk, whiteboard, and screen projector. A clicker was placed at the side of each computer. The tables were spaced far enough apart that discussions at one table did not disturb other groups.

The course instructor was a male senior chemistry teacher with significant Peer Instruction teaching experience, assisted by a technical assistant. He carried out the same learning session three times, once for each experimental condition. A few days before the 90-min chromatography lesson, the assistant divided the students randomly into three experimental conditions corresponding to different instruction sessions. The sessions took place one after another in the same classroom to avoid any communication between students in the different groups. They were organized in a fixed order, with the more innovative instruction session in the middle: Individual Instruction, Stepladder Peer Instruction, and Classic Peer Instruction. Ideally, the order of sessions should have been counterbalanced, but this was not possible due to the limited number of participants available.

On arrival, each student was given a designated seat in the classroom according to the first letter of their last name written on a Post-it® stuck on each computer and numbered from 1 to 4 (these numbers were only used for the Stepladder Peer Instruction group). In each group, the teacher informed the students that they had been randomly divided for practical reasons due to a limited number of computers. The students were instructed to use clickers to answer 10 multiple-choice questions during a tutorial class, which would give them formative feedback about their knowledge of chromatography

Table 1
Percentage of correct answers to the first test for each question in the pilot study and the present study.

Easy questions	Question #1	Question #2	Question #3	Question #4	Question #5	Total
Pilot study ($N = 50$)	80	52	66	74	52	64.8
Present study ($N = 84$)	63	58	57	64	41	56.6
Difficult questions	Question #6	Question #7	Question #8	Question #9	Question #10	
Pilot study ($N = 50$)	38	32	20	24	30	28.8
Present study ($N = 84$)	35	24	25	23	25	26.4

Table 2
Overview of the sequence for each instructional condition.

Instructional conditions			
Individual instruction	Classic peer instruction	Stepladder peer instruction	Time (total: 90 min)
First individual answer (Pre-test)			20 min.
		The first two group members discuss their answers	10 min.
		A third group member presents her/his ideas and solutions	5 min.
Students work individually to find answers to questions	Students discuss their answers in groups	Group discussion	10 min.
		A fourth group member presents her/his ideas and solutions	5 min.
		Group discussion	10 min.
Revised individual answer (Post-test) and whole class discussion			30 min.

acquired during previous lessons and from online resources containing course material, written documents, sound-slide presentations, videos, etc. To encourage the students to be fully involved, they were informed that the questions posed during this lesson would be helpful for self-evaluation before the exam. After a short training session using a few general knowledge questions, the experimental session was organized around a standard sequence (see Table 2). The course instructor introduced the questions one at a time in a consistent manner. The questions were given in the same order, and alternated according to their level of difficulty, easy or difficult. Students were given between one and three minutes (one minute for easy questions, and three minutes for difficult questions) to reflect silently and answer using the clickers without any access to (online) resources. They had to choose a response from three or four options, which included only one correct answer.

After each question, all the students received feedback in the form of a histogram displayed on the screen at the front of the classroom indicating the percentage of students who had selected each option. At this stage, the correct answer was not disclosed, and students were instructed to write down the percentage of responses for each option on a form with an empty grid with the questions numbered from 1 to 10 in rows, and the answers for each option from A to D in columns. They were also instructed to put a cross on the grid for their own answers. They were told that this grid would be used later to work on each question before revising their response. Depending on the instructional group to which they had been assigned, students were instructed to work either individually (Individual Instruction condition) or in four-member groups with different Peer Instruction procedures (Classic or Stepladder Peer Instruction).

In the *Individual Instruction* group, students were told to think individually about the 10 questions in any order for 40 min. During this period, the teacher did not communicate with the students and made sure that they did not communicate with each other. Students were given Internet access and they were encouraged to explore online resources to find the correct solutions. Obviously, the correct answers were not directly available and required thinking out. Students in all the groups were invited to use the form showing their own responses to the questions and the percentages of responses for each option of all the students.

In the two experimental Peer Instruction groups, students also had full Internet access and were also encouraged to explore their course notes and online resources.[3]

In the *Classic Peer Instruction* condition, students were instructed to think about all the questions and discuss them for 40 min with the students at the same table in order to find the correct answers. They were told that they would have a chance to revise their answer to each question at the end of their discussion and that no consensus was required, each student being free to give his/her own response irrespective of the responses of other group members. No other instructions were given.

In the *Stepladder Peer Instruction* group, participants were given a diagram describing the technique, as in previous studies (see Rogelberg et al., 1992). The diagram shows how the first two group members (#1 and #2) – the initial core group – remain together to discuss the questions for 10 min, while the other two members (#3 and #4) move into an adjacent classroom, waiting for a signal to join their group in turn.

While the two students of the initial core group discussed the questions of their choice in order to resolve their differences, the other students waited in an adjacent room and read scientific journals and newspapers unrelated to chromatography. They were instructed not to discuss the chromatography questions with the other students in the room. The assistant checked that they remained silent. In the learning context of the study, it would have been possible to allow the students to discuss their responses during this waiting period, but the time spent working on the questions would have been greater in the Stepladder PI than in the other conditions. Therefore, we chose not to give them this possibility, in order to remain in line with other studies on the Stepladder technique, which have mainly been conducted in laboratory settings. After 10 min, the

[3] Irrespective of the condition, it appeared that all the students explored online resources during the session, but because their behaviours were not recorded, these data cannot be quantified in the present study. To observe those behaviours, as well as other relevant behaviours such as the number and content of interactions between students, it would have been necessary to video each group, which was technically difficult in this learning context and would have exceeded the main objective of the present study.

teacher asked the next members (#3) to join their respective groups. They were given five minutes to read their own solutions to the series of questions, presenting them to the group and explaining their choice. After ten minutes of discussion in three-member groups, the last member (#4) joined the group, read and explain her/his solutions to the questions to the group during five minutes. At the end of the presentation, the whole group was given 10 min to discuss the questions; no consensus was required.

After the 40-min period, students used their clickers to give their revised individual solutions to the 10 multiple-choice chromatography questions, presented in the same order as the first time. Once all the participants had answered a question, the correct answer was displayed on the screen, and the class discussed it with their teacher for a varying amount of time, depending on the number of questions they posed and the amount of information provided by the teacher. The revised answers and the following whole class discussion lasted about 30 min. To conclude the learning session, students were asked to answer a single item questionnaire measuring their satisfaction with this mode of instruction compared to a traditional lecture.

2.4. Measures

2.4.1. Learning gain

Hake's (1998) gain statistic (g) was used as a standard metric in the present study to characterize an individual's improvement. To palliate certain negative effects of this measure, we also used an additional measure, the normalized change (c), derived from Hake's normalized gain (Marx & Cummings, 2007).

2.4.2. The normalized gain (g)

This measure is widely used in the Peer Instruction literature to evaluate the amount of knowledge learned by the student since the beginning of the session. It allows to compare increases in scores independently of pretest scores for different samples, and to avoid ceiling and floor effects. This measure is defined as the ratio of the average actual gain (% post − % pre) to the maximum possible average gain (100 − % pre): g = (%Post-test − %Pretest)/(100 − %Pretest), where %Pretest and %Post-test are the percentage of correct answers at the beginning and at the end of a session. For example, if a student gives 70% of correct answers at the Pretest and 90% at the Post-test, there is a 20% gain. The normalized gain is calculated by dividing this 20% by the possible amount of learning after the Pretest: 20/(100 − 70)% = 0.66. In this example, a gain of 0.66 indicates that 66% of the concepts to be learned were acquired at the end of the course. The numerator is the gross gain, while the denominator eliminates the bias due to differences of level at the beginning of the session. There are well-established norms to evaluate the magnitude of learning gain from the Hake factor (low gain: $g < 0.3$; medium gain: $0.7 < g > 0.3$; high gain: $g > 0.7$), whose scores range from $-\infty$ to $+1$.

2.4.3. The normalized change (c)

This is the ratio of the gain to the maximum possible gain, or of the loss to the maximum possible loss (Marx & Cummings, 2007). Normalized change can be used to eliminate the low pretest score bias and another disadvantage, which appears when some students' scores decrease between pre- and post-test. Normalized change c is similar to g when a post-test score is higher than the pretest score (most cases). By contrast, three other cases use different strategies to assign a c value: (1) students who score either 0 or 100 on both pre- and post-tests are removed from the data set, because otherwise the c-value would be recorded as 0; (2) students who have identical pre- and post-test scores are assigned a c-value equal to zero; and (3) for students who perform worse on the post-test than on the pretest, the following formula is used: c = (%Post-test − %Pretest)/%Pretest. Using this strategy, normalized change scores range from −1 to +1, generating a symmetric range of scores.

2.4.4. Satisfaction

Immediately after the session, students had to complete a single-item questionnaire about the learning session in comparison to a lecture (“I prefer this type of instruction to a traditional lecture”) on a four-point Likert scale (from 1 = *not at all* to 4 = *absolutely*).

3. Results

3.1. Learning gain

The statistical analyses compared normalized gains (g) across the instructional conditions according to the level of difficulty of the questions. A 3 X 2 mixed ANCOVA, with Instructional condition as the between-subjects factor (Individual Instruction vs. Classic Peer Instruction vs. Stepladder Peer Instruction), and Difficulty of questions (Easy vs. Difficult) as the within-subjects factor, was conducted to examine the effects on learning gains. Gender was treated as a covariate because of its potential influence on learning in Peer Instruction (e.g., Jones et al., 2012; King & Joshi, 2008; Lorenzo, Crouch, & Mazur, 2006).

Statistical analyses did not reveal an effect of Gender, nor a significant main effect of Difficulty on learning gains (both $Fs < 1.0$). The learning gain for easy questions did not differ from that for difficult questions, both remaining at a medium level (Ms = 0.47 and 0.52, respectively). As expected in Hypothesis 1, a main effect of Instructional method was observed on

Table 3
Mean percentages of correct answers on Pre- and Post-test, normalized learning gains (Standard Deviations in parentheses) on easy and difficult questions for each instructional condition.

Instructional condition	Easy questions			Difficult questions		
	Pre-test (%)	Post-test (%)	Normalized gain (g)	Pre-test (%)	Post-test (%)	Normalized gain (g)
Individual instruction (control) ($N = 28$)	64.28	85.71	0.49 (0.48)	29.28	49.28	0.23 (0.33)
Classic Peer Instruction ($N = 28$)	69.28	90.71	0.43 (0.71)	26.43	68.57	0.53 (0.31)
Stepladder Peer Instruction ($N = 28$)	68.57	87.85	0.49 (0.49)	38.57	88.57	0.80 (0.29)
Total	67.38	88.57	0.47 (0.57)	31.43	68.81	0.52 (0.39)

learning gains, $F(2, 80) = 6.015$, $p = .004$, partial $\eta^2 = 0.13$ (see Table 3). The learning gains decreased linearly over the conditions: Stepladder Peer Instruction, Classic Peer Instruction, and Individual Instruction (*Ms* = 0.65, 0.49, 0.36, respectively).

More interesting, and as expected in Hypothesis 2, this effect was qualified by an interaction effect between Instructional condition and Difficulty of questions on learning gains, $F(2, 80) = 4.614$, $p = .013$, partial $\eta^2 = 0.10$. As Fig. 1 shows, there were no significant differences between the three instructional conditions for the easy questions, while differences between conditions were observed for difficult questions. Planned contrasts were performed on learning gains for difficult questions. The first compared the control condition of Individual Instruction without peer discussion (−2) with the two experimental conditions where discussions with peers were encouraged, namely, Stepladder Peer Instruction (1) and Classic Peer Instruction (1). Results revealed that learning gains were greater under the two Peer Instruction conditions than under the control condition, $F(1, 83) = 31.421$, $p = .001$, partial $\eta^2 = 0.28$ (*Ms* = 0.66 and 0.23, respectively). The second planned contrast compared the two experimental conditions, Stepladder Peer Instruction (−1) and Classic Peer Instruction (1); in this analysis, the control condition was coded 0. Results revealed that learning gains were greater with Stepladder Peer Instruction than with Classic Peer Instruction, $F(2, 83) = 22.556$, $p = .001$, partial $\eta^2 = 0.36$ (*Ms* = 0.80 and 0.53, respectively). These results show that peer interaction improves learning gains. More importantly, they also indicate that the Stepladder technique improves the learning gains produced by Peer Instruction in comparison to the more conventional method where students join groups simultaneously.

Similar statistical analyses were performed on the normalized change (c) in order to eliminate the low pretest score bias, and another disadvantage, which appears when some students' scores decrease between pre- and post-test (see Marx & Cummings, 2007). In the present sample, we observed four such cases for easy questions and six cases for difficult questions, justifying this calculation method (8.4%). As the results obtained with this additional measure were identical to those observed with normalized gain (g), they are presented separately in a footnote.[4]

Finally, additional descriptive analyses were performed to determine how many students in each instructional group gave correct answers to all questions (i.e. 100% of correct responses). As Table 4 shows, the greatest benefit occurred in the Stepladder Peer Instruction group, in which the number of students giving correct answers to all questions jumped from 1 to 16 (out of 28 students) for the five difficult questions compared to 1 to 4 in the Classic Peer Instruction group. In the Individual instruction group, none of the students found all the correct answers to the difficult questions, either at pretest, or at post-test. Unsurprisingly, a similar improvement between Pre- and Post-test was found in all instructional groups for the easy questions. The results show that 51.19% of students found the correct answers to the five easy questions at the end of the session, against only 23.80% for the five difficult questions (more than double that number in the Stepladder PI condition, 57.14%). These findings indicate that just over half the students improved their performance, giving zero incorrect responses to the easy questions, and only a quarter of the students gave zero incorrect responses to the difficult questions, suggesting a margin of improvement.

3.2. Satisfaction

An ANCOVA with Instructional condition as the between-subjects factor (Individual Instruction vs. Classic Peer Instruction vs. Stepladder Peer Instruction), and Gender as covariate, was performed on perceived satisfaction. As previous studies showed that the Stepladder technique yielded better satisfaction than discussion in conventional groups, there was no justification for planned contrasts on this measure. A significant effect of Instructional condition was observed, whereby students were less satisfied with Individual Instruction ($M = 2.10$ and $SD = 0.86$) than with Classic Peer Instruction ($M = 2.55$ and $SD = 0.93$) and Stepladder Peer Instruction ($M = 2.70$ and $SD = 0.92$), $F(1, 82) = 3.463$, $p < .04$, partial $\eta^2 = 0.08$. Pairwise

[4] An ANCOVA yielded a main effect of Instructional method on learning gains measured with the normalized change metric, $F(2, 65) = 7.95$, $p = .001$, partial $\eta^2 = 0.19$. The learning gains decreased linearly across conditions: Stepladder Peer Instruction, Classic Peer Instruction, and Individual Instruction (*Ms* = 0.68, 0.56, 0.34, respectively). This effect was qualified by an interaction effect between Instructional condition and Difficulty of questions on learning gains, $F(2, 65) = 4.87$, $p = .011$, partial $\eta^2 = 0.13$, revealing no significant differences between the three instructional conditions for the easy questions ($M_{\text{Stepladder PI}} = 0.59$, $M_{\text{Classic PI}} = 0.63$, and $M_{\text{Individual Instruction}} = 0.47$), while differences between conditions were observed for difficult questions, the Stepladder Peer Instruction condition producing better performance ($M_{\text{Stepladder PI}} = 0.83$, $M_{\text{Classic PI}} = 0.51$, and $M_{\text{Individual Instruction}} = 0.20$).

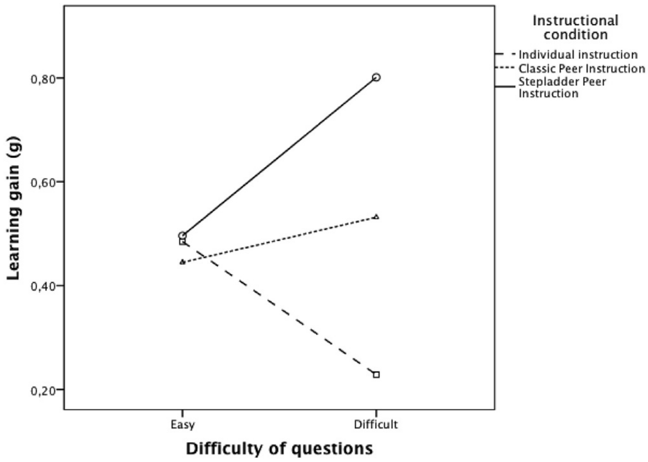


Fig. 1. Effect of instructional condition (Individual Instruction, Classic Peer instruction or Stepladder Peer Instruction) and difficulty of multiple-choice questions (Easy and Difficult) on learning gains (g).

comparisons revealed that Individual Instruction differed from both Classic Peer Instruction ($p < .05$) and Stepladder Peer Instruction ($p < .01$). The two Peer Instruction conditions did not significantly differ.

4. Discussion

The goal of the present study was to improve Peer Instruction, and more specifically the learning gains generated by this method in fundamental sciences, by using the Stepladder technique. As described above, this is a technique derived from applied psychology research, which is known to be effective in boosting performance, at least in laboratory settings, when team building simulation tasks are used (e.g., Orpen, 1997; Rogelberg et al., 1992; Rogelberg & O'Connor, 1998). In higher education settings where the Peer Instruction method is used to improve learning in fundamental sciences, it allows all the students, and not just some of them, to engage actively in discussions during a PI session. This optimization of the PI method should also contribute to a greater increase in learning gains. Indeed, examination of the data gathered in PI studies reveals that learning gains generally remain at a medium level (i.e. around a value of 0.5), and a difficulty to exceed this level is generally observed (e.g., Crouch & Mazur, 2001; Fagen et al., 2002; Freeman et al., 2014; Hake, 1998). One of the reasons was thought to be that students in groups do not all engage actively in fruitful discussions with their peers during a PI session and hence do not all benefit from social interaction to improve their learning. Based on this reasoning, it was expected that the Stepladder technique, in which the entry of members to a group is structured so that every member is forced to participate during a Peer Instruction session, might yield greater learning gains than classic PI and individual instruction without any discussion with peers.

As expected, our results showed that learning gains were greatest in the Stepladder PI group, and, maybe more importantly, this effect was mainly observed for difficult questions. The findings provide the first empirical evidence that Peer Instruction can produce higher learning gains than previously demonstrated, exceeding the medium gain generally found in the literature and attaining a high gain when combined with the Stepladder technique, which forces all the students to participate in turn in discussions. These effects were found irrespective of the method used to measure learning gains, either normalized gain (Hake, 1998) or normalized change (Marx & Cummings, 2007).

Apart from the effect on learning, our results also show that perceived satisfaction was higher when students had been engaged in peer discussions, suggesting the affective benefits of social interactions between peers. These findings are consistent with results of a recent study (Morice et al., in press) demonstrating that students perceive the Peer Instruction method as more satisfying, engaging, and useful than an individual learning method where they have to work alone. However, in contrast to laboratory experiments in which individuals in groups using the Stepladder technique were highly satisfied with this structured group technique (Rogelberg et al., 1992), the present study failed to produce statistical evidence

Table 4
Number of students providing **all** correct answers to easy and difficult questions at Pre- and Post-test for each instructional condition (percentages given in parentheses).

Instructional condition	Easy questions		Difficult questions	
	Pre-test	Post-test	Pre-test	Post-test
Individual instruction (control) ($N = 28$)	5 (17.86%)	12 (42.86%)	0 (0.0%)	0 (0.0%)
Classic Peer Instruction ($N = 28$)	6 (21.43%)	18 (64.28%)	1 (3.57%)	4 (12.28%)
Stepladder Peer Instruction ($N = 28$)	6 (21.43%)	13 (46.43%)	1 (3.57%)	16 (57.14%)
Total ($N = 84$)	17 (20.23%)	43 (51.19%)	2 (2.38%)	20 (23.80%)

of greater satisfaction with the PI method linked to the Stepladder technique than when the same method was used without the consecutive introduction of group members. This discrepancy could be explained by the time separating the discussions between peers, and the completion of the single-item questionnaire after the 30-min whole-class discussion. The preliminary discussions under the PI conditions may have prepared students for a larger discussion with the whole class, leading them to express greater satisfaction than under the Individual Instruction condition. Unfortunately, the present study did not include measures to support this explanation. Another limitation of the present study is that discussions with peers may have been biased towards the most common answer chosen by the class (or by a group member) in the Peer Instruction conditions, exhibiting a conformity effect (e.g., Asch, 1951; Deutsch & Gerard, 1955). Indeed, in a study of students' responses during Peer Instruction in two chemical thermodynamics classes, Brooks and Koretsky (2011) showed that the majority of students selected the consensual answer after group discussion, irrespective of whether it was right or wrong. Another study found that this effect only appeared when students could see the response graph (Perez et al., 2010); students who could see the bar graph were found to be 30% more likely to switch from a less common to the most common response, although this effect was less pronounced for multiple-choice than for true/false questions. In the present study, we believe that a conformity effect is unlikely, due to the systematic increase in the proportion of correct responses from pre-to post-test. Moreover, the number of students providing all correct answers was very similar in the three conditions for easy questions, but higher in the Stepladder PI than in the other two conditions for difficult questions. As a majority influence should appear on both easy and difficult questions, including incorrect answers, it seems that no such process was at work during the PI session. Nevertheless, future research should examine this possibility more thoroughly, by displaying (or not) the histogram of responses to the participants, placing them in a more or less strong uncertainty situation.

As far as we know, this is the first experiment in the field of higher education demonstrating that the Peer Instruction method can be enhanced by a structured technique known to optimize students' participation in group discussions, and consequently increase learning gains. While collaborative learning based on Peer Instruction may yield better performance than individual work, the way the social interactions are structured in collaborating groups is also important. Organizing students in groups does not automatically mean that they will participate effectively in group discussions, and teacher guidance is often required to improve collaborative learning (e.g., Corden, 2001; Kirschner, Sweller, & Clark, 2006). This guidance may be 'instrumental', i.e. the teacher organizes the learning task as a group activity, but it may also be 'socio-relational', whereby the teacher guides the groups to stimulate social interactions for a more efficient collaboration in learning tasks. Based on this 'socio-relational' role, the Stepladder technique may offer teachers a useful structured method to help students conduct their group discussions, inviting them to explain their ideas and solutions in turn. This type of guidance is consistent with a number of studies investigating how collaborative learning can be fostered by using 'collaboration scripts' to structure social interaction among students (e.g., Kobbe et al., 2007; Weinberger et al., 2002). To our knowledge, none of these studies have mentioned the Stepladder technique as a useful and scripted method for improving the efficacy of Peer Instruction in the fundamental sciences. Consequently, the present findings suggest a new teacher guidance technique, which, like 'collaboration scripts', is based on the management of social interactions in groups, engaging all the students in a collaborative learning process by stimulating peer discussions. Because the Stepladder technique may improve the efficacy of the Peer Instruction often used in STEM disciplines (science, technology, engineering, and mathematics), it is reasonable to consider that it may be an effective and innovative teaching tool to make (good) scientists (e.g., Drane, Micari, & Light, 2014; Light & Micari, 2013).

5. Conclusion

The present research suggests that the Stepladder technique may be fruitfully extended to higher education, and may help further increase learning gains among students involved in a Peer Instruction session. While a large number of studies have compared the effectiveness of PI with traditional lectures (e.g., Hake, 1998; Mayer et al., 2009), the present study compared different ways of designing a PI session. As such, it may be considered as a paradigm shift from 'first-generation research', which has mainly compared traditional lecturing with active learning methods, to 'second-generation research' "*using advances in educational psychology and cognitive science to inspire changes in course design*" (Freeman et al., 2014, p. 8413).

Finally, the present study may be considered as a first attempt to compare instructional methods where students have to interact with their peers, simultaneously or sequentially, with an individual instruction method without any social interaction. Although the present study is limited in scope, it offers a further step in the Peer Instruction literature, showing how learning gains can be enhanced when students are instructed to argue their points of view one by one instead of participating simultaneously in group or class discussions. As few studies have so far examined Peer Instruction in educational psychology, we hope that the present study will arouse interest among researchers in this field of research, helping them to design fruitful instructional methods to improve learning among higher education students, not only in the STEM disciplines, but also in many other disciplines in which multiple-choice questions are used.

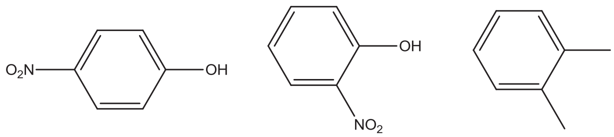
Acknowledgements

This research was supported by the Maison des Sciences de l'Homme de Bretagne (Britanny, France). Special thanks to Master degree students for their help in collecting and coding data. We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions.

Appendix

Easy question

In silica gel column chromatography, in which order would the following compounds be eluted?



1 2 3

A. 3 then 1 then 2.
B. 3 then 2 then 1.
C. 1 then 2 then 3.

Difficult question

Using the chromatographic data shown in the table (A = barbitone, B = amobarbitone, C = secobarbitone), calculate the response factors of amobarbitone and secobarbitone with respect to barbitone, and hence their amount in the test sample.

Retention time (min)	Area (standard)	Area (%)	Retention time (min)	Area (sample)	Area (%)
0.26	7,335,600	97.315	0.24	9,512,000	98.186
2.41	105,020	1.393	2.41	97,273	1.004
4.91	54,984	0.729	4.49	39,295	0.406
6.68	42,408	0.563	6.82	39,187	0.405

A. mB = 38.6 mg, mC = 49.9 mg.
B. mB = 49.9 mg, mC = 38.9 mg.
C. mB = 39.3 mg, mC = 38.2 mg.

References

Asch, S. E. (1951). Effects of group pressure on the modification and distortion of judgements. In H. Guetzkow (Ed.), *Groups, leadership and men* (pp. 177–190). Pittsburgh, PA: Carnegie Press.

Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: blending, flipping, and efficiency in active learning classrooms. *Computers and Education, 78*, 227–236. http://dx.doi.org/10.1016/j.compedu.2014.06.006.

Banks, D. (2006). *Audience response systems in higher education: Applications and cases*. Hershey, PA: Information Science Publishing.

Brooks, B. J., & Koretsky, M. D. (2011). The influence of group discussion on students' responses and confidence during Peer Instruction. *Journal of Chemical Education, 88*(11), 1477–1484, 10.1021/ed101066x.

Corden, R. E. (2001). Group discussion and the importance of a shared perspective: learning from collaborative research. *Qualitative Research, 1*(3), 347–367. http://dx.doi.org/10.1177/146879410100100305.

Crouch, C. H., & Mazur, E. (2001). Peer Instruction: ten years of experience and results. *American Journal of Physics, 69*(9), 970–977. http://dx.doi.org/10.1119/1.1374249.

Crouch, C. H., Watkins, J., Fagen, A. P., & Mazur, E. (2007). Peer Instruction: engaging students one-on-one, all at once. *Reviews in Physics Education Research, 1*(1), 40–95.

Dancy, M., & Henderson, C. (2010). Pedagogical practices and instructional change of physics Faculty. *American Journal of Physics, 78*(10), 1056–1063. http://dx.doi.org/10.1119/1.3446763.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology, 51*, 629–636. http://dx.doi.org/10.1037/h0046408.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). *The evolution of research on collaborative learning. Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189–211). Oxford: Elsevier.

Doise, W., & Mugny, G. (1984). *The social development of the intellect*. Oxford: Pergamon Press.

Drane, D., Micari, M., & Light, G. (2014). Students as teachers: effectiveness of a peer-led STEM learning program over 10 years. *Educational Research and Evaluation, 20*, 210–230. http://dx.doi.org/10.1080/13803611.2014.895388.

Fagen, A. P., Crouch, C. H., & Mazur, E. (2002). Peer Instruction: results from a range of classrooms. *Physics Teacher, 40*, 206–209. http://dx.doi.org/10.1119/1.1474140.

Fantuzzo, J. W., Riggio, R. E., Connelly, S., & Dimeff, L. A. (1989). Effects of reciprocal peer tutoring on academic achievement and psychological adjustment: a component analysis. *Journal of Educational Psychology, 81*, 173–177. http://dx.doi.org/10.1037/0022-0663.81.2.173.

Freeman, S., Eddy, S., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415. http://dx.doi.org/10.1073/pnas.1319030111.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*, 265–348. http://dx.doi.org/10.1016/j.learninstruc.2009.08.007.

Goodyear, P., Jones, C., & Thompson, K. (2014). Computer-supported collaborative learning: instructional approaches, group processes, and educational designs. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 439–451). New york, NY: Springer.

Hake, R. (1998). Interactive engagement versus traditional methods: a six-thousand student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74. http://dx.doi.org/10.1119/1.18809.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*, 141–158. http://dx.doi.org/10.1119/1.2343497.

Hockings, S. C., DeAngelis, K. J., & Frey, R. F. (2008). Peer-led team learning in general chemistry: implementation and evaluation. *Journal of Chemical Education, 85*, 990–996, 10.1021/ed075p.185.

James, M. C. (2006). The effect of grading incentive on student discourse in Peer Instruction. *American Journal of Physics, 74*, 689–691. http://dx.doi.org/10.1119/1.2198887.

Johnson, D., & Johnson, R. (2008). Cooperation and the use of technology. In J. M. Spector, M. D. Merrill, J. van Merrienboer, & M. Driscoll (Eds.), *Handbook of research on educational communications and technology* (3rd ed., pp. 1017–1044). New York, NY: Taylor & Francis.

Jones, M. E., Antonenko, P. D., & Greenwood, C. M. (2012). The impact of collaborative and individualized student response system strategies on learner motivation, metacognition, and knowledge transfer. *Journal of Computer Assisted Learning, 28*(5), 477–487. http://dx.doi.org/10.1111/j.1365-2729.2011.00470.x.

King, D. B., & Joshi, S. (2008). Gender differences in the use and effectiveness of personal response devices. *Journal of Science Education and Technology, 17*(6), 544–552. http://dx.doi.org/10.1007/s10956-008-9121-7.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychology, 41*(2), 75–86. http://dx.doi.org/10.1207/s15326985ep4102_1.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., & Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*, 211–224. http://dx.doi.org/10.1007/s11412-007-9014-4.

Lantz, M. E. (2010). The use of "Clickers" in the classroom: teaching innovation or merely an amusing novelty? *Computers in Human Behavior, 26*(4), 556–561. http://dx.doi.org/10.1016/j.chb.2010.02.014.

Lasry, N., Charles, E., Whittaker, C., & Lautman, M. (2009). When talking is better than staying quiet. In M. Sabela, C. Henderson, & C. Singh (Eds.), *Proceedings from the physics education research conference* (pp. 181–184). College Park, MD: American Institute of Physics.

Lasry, N., Mazur, E., & Watkins, J. (2008). Peer Instruction: from Harvard to community colleges. *American Journal of Physics, 76*(11), 1066–1069. http://dx.doi.org/10.1119/1.2978182.

Lewis, S., & Lewis, J. E. (2005). Departing from lectures: a Peer-led guided inquiry alternative. *Journal of Chemical Education, 82*, 135–139, 10.1021/ed082p.135.

Light, G., & Micari, M. (2013). *Making scientists: Six principles for effective college teaching*. Cambridge: Harvard University Press.

Lorenzo, M., Crouch, C. H., & Mazur, E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics, 74*(2), 118–122. http://dx.doi.org/10.1119/1.2162549.

Lou, Y., Abrami, P., & d'Appolonia, S. (2001). Small group and individual learning with technology: a meta-analysis. *Review of Educational Research, 71*, 449–521. http://dx.doi.org/10.3102/00346543071003449.

Lucas, A. (2009). Using Peer Instruction and i-clickers to enhance student participation in calculus. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, 19*, 219–231. http://dx.doi.org/10.1080/10511970701643970.

Marx, J., & Cummings, K. (2007). Normalized change. *American Journal of Physics, 75*, 87–91. http://dx.doi.org/10.1119/1.2372468.

Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., et al. (2009). Clickers in college classrooms: fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology, 34*(1), 51–57. http://dx.doi.org/10.1016/j.cedpsych.2008.04.002.

Mazur, E. (1997). *Peer instruction: A user's manual series in educational innovation*. Upper Saddle River, NJ: Prentice Hall.

McCroskey, J. C., & Beatty, M. J. (1984). Communication apprehension and accumulated communication state anxiety experiences: a research note. *Communication Monographs, 51*, 79–84. http://dx.doi.org/10.1080/03637758409390185.

Micari, M., & Pazos, P. (2014). Worrying about what others think: a social-comparison concern intervention in small learning groups. *Active Learning in Higher Education, 15*(3), 249–262. http://dx.doi.org/10.1177/1469787414544874.

Morice J., Michinov N., Delaval M., Sideridou A., and Ferrières V., Comparing the effectiveness of Peer Instruction to individual learning during a chromatography course in a blended learning environment, *Journal of Computer Assisted Learning*, (in press).

Orpen, C. (1997). Using the Stepladder technique to improve team performance. *Psychological Studies, 42*(1), 24–28. http://dx.doi.org/10.1108/13527599510084830.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*, 117–175. http://dx.doi.org/10.1207/s1532690xci0102_1.

Perez, K. E., Strauss, E. A., Downey, N., Galbraith, A., Jeanne, R., & Cooper, S. (2010). Does displaying the class results affect student discussion during peer instruction? *Cell Biology Education-A Journal of Life Science Education, 9*, 133–140. http://dx.doi.org/10.1187/cbe.09-11-0080.

Perlmutter, M., Behrend, S. D., Kuo, F., & Muller, A. (1989). Social influences on children's problem solving. *Developmental Psychology, 25*, 744–754. http://dx.doi.org/10.1037/0012-1649.25.5.744.

Porter, L., Bailey-Lee, C., Simon, B., & Zingaro, D. (2011). Peer Instruction: do students really learn from peer discussion in computing?. In *Proceedings of the seventh international workshop on computing education research* (pp. 45–52). ACM.

Rao, S. P., & DiCarlo, S. E. (2000). Peer Instruction improves performance on quizzes. *Advances in Physiology Education, 24*(1), 51–55.

Rogelberg, S. G., Barnes-Farrell, J. L., & Lowe, C. A. (1992). The Stepladder technique: an alternative group structure facilitating effective group decision making. *Journal of Applied Psychology, 77*, 730–737. http://dx.doi.org/10.1037//0021-9010.77.5.730.

Rogelberg, S. G., & O'Connor, M. S. (1998). Extending the Stepladder technique: an examination of self-paced Stepladder groups. *Group Dynamics: Theory, Research and Practice, 2*, 82–91. http://dx.doi.org/10.1037/1089-2699.2.2.82.

Rogelberg, S. G., O'Connor, M. S., & Sederburg, M. (2002). Using the Stepladder technique to facilitate the performance of audio conferencing groups. *Journal of Applied Psychology, 87*, 994–1000. http://dx.doi.org/10.1037/0021-9010.87.5.994.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: a review of the research. *Review of Educational Research, 64*, 479–530. http://dx.doi.org/10.3102/00346543064004479.

Singh, C. (2005). Impact of peer interaction on conceptual test performance. *American Journal of Physics, 73*, 446–451. http://dx.doi.org/10.1119/1.1858450.

Slavin, R. E. (2011). Instruction based on cooperative learning. In R. E. Mayer, & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 344–360). New York: Taylor & Francis.

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. http://dx.doi.org/10.1126/science.1165919.

Springer, L., Stanne, M., & Donovan, S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. *Review of Educational Research, 69*, 21–51. http://dx.doi.org/10.3102/00346543069001021.

Thompson, L. F., & Coovert, M. D. (2002). Stepping up to the challenge: a critical examination of face-to-face and computer-mediated team decision making. *Group Dynamics: Theory, Research, and Practice, 6*, 52–64. http://dx.doi.org/10.1037/1089-2699.6.1.52.

Van Zundert, M., Sluijsmans, D., & Van Merriënboer, J. (2010). Effective peer assessment processes: research findings and future directions. *Learning and Instruction, 20*, 270–279. http://dx.doi.org/10.1016/j.learninstruc.2009.08.004.

Wamser, C. C. (2006). Peer-led team learning in organic chemistry: effects on student performance, success, and persistence in the course. *Journal of Chemical Education, 83*, 1562–1566, 10.1021/ed083p.1562.

Watkins, J., & Mazur, E. (2010). Just-in-Time teaching and peer instruction. In S. Simkins, & M. H. Maier (Eds.), *Just in time teaching: Across the disciplines, and across the academy* (pp. 39–62). Sterling, VA: Stylus Publishing.

Weber, K., Maher, C., Powell, A., & Lee, H. S. (2008). Learning opportunities from group discussions: warrants become the objects of debate. *Educational Studies in Mathematics, 68*(3), 247–261. http://dx.doi.org/10.1007/s10649-008-9114-8.

Weinberger, A., Fischer, F., & Mandl, H. (2002). Fostering computer-supported collaboration with cooperation scripts and scaffolds. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations of a CSCL community* (pp. 573–574). Mahwah, NJ: Erlbaum.

Winquist, J. R., & Franz, T. M. (2008). Does the Stepladder technique improve group decision making? A series of failed replications. *Group Dynamics: Theory, Research, and Practice, 12*, 255–267. http://dx.doi.org/10.1037/1089-2699.12.4.255.

Zingaro, D., & Porter, L. (2014). Peer Instruction in computing: the value of instructor intervention. *Computers and Education, 71*, 87–96. http://dx.doi.org/10.1016/j.compedu.2013.09.015.