Collaborative Group Learning in a Swiss introductory physics class

G Feldman$^{1,2}$, G Schiltz$^2$ and A Vaterlaus$^2$

$^1$ George Washington University, Department of Physics, Washington DC, USA
$^2$ ETH Zürich, Department of Physics, Zürich, Switzerland

feldman@gwu.edu

Abstract. We report on the implementation of the SCALE-UP collaborative pedagogical approach in a first-year undergraduate physics class at ETH Zürich. The performance of the students on various assessments was compared with identical assessments in a parallel lecture class in order to gauge the effectiveness of this approach. Surveys were also given to get feedback on how students were reacting to this innovative style of pedagogy.

1. Introduction

Studies of undergraduate STEM education have demonstrated that students need to be actively engaged in the learning process in order for it to be effective. A passive lecture environment (“teaching by telling”) has been shown to be largely ineffective in developing students’ skills in critical thinking and problem solving [1]. Despite the abundant data which indicate the shortcomings of the conventional lecture format, however, it continues to be the instructional mode that is prevalent in most institutions.

In the United States, active-learning strategies are becoming more common in physics classes, although most European classes still rely on the standard lecture approach. One active-learning technique that has gained popularity in the U.S. is the SCALE-UP (or studio physics) group-learning approach that was developed at North Carolina State University [2,3] and has been adopted by over 200 institutions around the U.S. (including George Washington University). In this classroom format, students work together in small groups on guided activities and the instructor serves as more of a facilitator or “coach” instead of a lecturer. When one of us (G. Feldman) had a sabbatical semester in Spring 2017, we embarked on a pilot implementation of SCALE-UP in an introductory physics class at ETH Zürich. The project included a preliminary 6-week visit to ETH in August/September 2016 for planning and preparation, followed by the delivery of the SCALE-UP class itself in the 14-week spring semester (February–June 2017).

2. Classroom Setup

The Physics 1 introductory class at ETH Zürich consisted of approximately 370 students, in a range of disciplines including environmental science, food science, earth science and agricultural science. Out of this student population, we offered the SCALE-UP collaborative class to 54 students on a voluntary enrolment basis. Registration for the class took place in December 2016 and there were 94 students...
Figure 1. Classroom configuration with 9 hexagonal tables for SCALE-UP (right panel) as adapted from a standard classroom layout with movable trapezoidal tables (left panel).

Figure 2. Photos of the SCALE-UP classroom (top left) and the lecture hall (top right). The two photos on the left show students working on their whiteboards, assisted by the instructor and the TAs. The photo on the bottom right shows a student presenting his group’s results on their whiteboard.

interested in signing up – as a result, we found ourselves with a waiting list of 40 students. The remaining students (approximately 320 of them) constituted the population in the parallel lecture class that was compared with the SCALE-UP class.

The classroom was configured by using 9 hexagonal tables (formed with two trapezoidal tables) to organize the 54 students into 18 groups of three (two groups per table). The initial classroom layout and the reconfigured SCALE-UP format are shown in Fig. 1. Portable whiteboards (60 cm × 90 cm) were distributed to each group for working together on numerical and conceptual questions in class. An electronic response system (“clickers”) based on TurningPoint [4] was used to deliver conceptual questions to the class, following the Peer Instruction method pioneered by Mazur [5]. Some small
demonstration equipment and an instructor tablet computer were provided by the Department of Physics at ETH to support the classroom activities. Several photos of the classroom itself, as compared to the lecture hall, are depicted in Fig. 2.

While the large lecture class (with A. Vaterlaus) was delivered in German, the SCALE-UP class (with G. Feldman) was taught in English. This did not pose any particular problem for the Swiss students, since their English language abilities were uniformly excellent (even for first-year students). To facilitate pre-class preparation and post-class problem-solving practice, we obtained 60 student licenses from Pearson publishers for the MasteringPhysics online system [6], with a link to the Giancoli introductory physics textbook [7]. This served as the primary textbook for the SCALE-UP class, and conceptual and numerical exercises were offered online for the students. The acquisition of these 60 licenses was fully funded by the Department of Physics at ETH, so the students received this benefit for free.

The instructional schedule for both settings consisted of one 3-hr class on Monday morning and a 1-hr recitation session on Thursday with a Teaching Assistant (TA). Since each academic hour is actually 50 minutes followed by a 10-minute break, the in-class contact time essentially amounted to only 150 mins = 2.5 hrs. The SCALE-UP class had three TA’s who assisted with the collaborative activities in the classroom on Monday (with G. Feldman) and who ran the recitation sessions (each with about 18 students) which mainly served as question/answer periods. The Physics 1 class did not have an associated laboratory session.

The physics topics covered in both classes were totally aligned and were typical for a first-semester physics course: kinematics, forces, energy, momentum, rigid-body motion, torque, oscillations, fluids, heat and thermodynamics. Students were fully expected to come to class prepared to work on the group activities together, which included conceptual questions that they would discuss and numerical problems that they would solve. Solutions to these questions and problems were reported by the students themselves in the classroom, displaying their group’s work on their whiteboards to the rest of the class and explaining their steps in the solution. Formal lecture in the conventional sense was reduced to a minimal level, and while there were certainly many engaging discussions about the material (either with the TA’s in the groups as they worked or by the instructor with the entire class as solutions were presented), there was no extensive “transmission” of vast amounts of physics content for extended periods of time. The focus was primarily on the students and allowing them to wrestle with the material themselves (with real-time guidance by the instructors, as needed) – it was truly a “student-centered” classroom.

3. Pedagogical Methodology

It is clear that there were quite a few novel features in the SCALE-UP class that were very different from the lecture environment to which the students were accustomed. Such innovative pedagogical strategies as collaborative activities with the portable whiteboards, the use of “clickers” to answer conceptual questions [4,5], and the availability of the online questions/problems delivered in the MasteringPhysics system [6] – these were all new and unfamiliar to the students.

As research questions, we wanted to explore how the students would respond to these features, as well as other considerations that applied uniquely to the SCALE-UP approach:

- no lectures – very little formal “transmission” of information
- necessity of coming to class prepared to work on exercises
- reactions to new class elements (clickers, whiteboards, online homework, etc.)
- willingness to speak up and report solutions to the entire class
- different (and dynamic!) classroom environment compared to lecture

What did the students actually do in the SCALE-UP class? The in-class activities consisted of “ponderables” (conceptual or numerical exercises to think about) and “tangibles” (short hands-on exercises or demos involving some equipment and some manipulation). Responses to conceptual questions were recorded using TurningPoint clickers [4] and displayed to the class for discussion and
resolution. Some graphical or numerical problems were solved collectively by the group members using their whiteboards and then reported to the entire class. Tangibles included short activities such as: (1) dropping a meterstick and catching it to measure human reaction time via free-fall acceleration, (2) balancing a meterstick on a fulcrum (like a see-saw) with a known mass on one side in order to find the unknown mass of the meterstick, (3) submerging a weight into a beaker of water on a scale to determine if the scale reading changes due to the buoyant force on the weight, and (4) playing with Newton’s Cradle to observe elastic collisions between steel balls.

Outside of class, all the Physics 1 students (SCALE-UP and lecture) had homework problems assigned in the Moodle online system used at ETH. The weekly problem sets consisted of about 4-5 numerical problems that were optional and were intended to be discussed with the TA’s in the recitation sessions. These assignments could be handed in for feedback from the TA, but they were not formally graded (and most students did not hand them in anyway).

For the SCALE-UP class, the use of the MasteringPhysics online system [6] enabled the creation of pre-class “Warmups” which were conceptual questions designed to help the students gauge their own level of understanding before coming to class. Reading the Giancoli textbook in advance (available as an e-text in MasteringPhysics) and working on the Warmups were the primary means of preparation available to the SCALE-UP students. It is important to note that this level of pre-class work was fully expected of the students, and their responsibility for preparing before class was made abundantly clear to them. In addition, online homework problems were also offered after class to the students through MasteringPhysics to provide more opportunities for numerical problem-solving practice. While the Warmups and homework sets served as a valuable resource for the students throughout the semester, it is necessary to point out that all of the exercises in MasteringPhysics were purely optional. It should also be noted that since we had only obtained 60 student licenses for this product, the online Warmups and homework sets were exclusively available to the students in the SCALE-UP class.

To address our research questions and to make some determination about the effectiveness of the SCALE-UP pedagogy compared to the lecture format, we collected data (see Table 1) to provide both a quantitative evaluation and some qualitative feedback on the general impact of our intervention. For summative assessments, we gave an optional mid-term exam to both classes in the tenth week of the semester, and we will give a required common Final Exam (for Physics 1+2 combined) in mid-January of 2018. As an incentive to take the mid-term exam, that score could be counted for 10% of the final Physics 1+2 grade, such that it was a “bonus” that could partially compensate for the Final Exam. Note that there was no comprehensive exam at the end of Physics 1. For a diagnostic conceptual assessment, we gave the Force Concept Inventory (FCI) [8,9] as a pre-test in the first week of class and as a post-test on the last day of class (both were optional). For formative assessments, we gave short (15 minutes) conceptual/numerical quizzes every three weeks in the recitation sessions (which were also optional). Finally, for student feedback, we offered a mid-semester questionnaire for the SCALE-UP class to check how things were going, and there was also a standard ETH course evaluation administered at the end of the semester (for all courses across the institution).

4. Results
The mid-term exam was given to both classes in the tenth week of the semester, and the results are shown in Fig. 3 below. The mid-term consisted of three conceptual questions and three numerical questions (each worth 12 points), for a total of 72 points. In the figure, the scores are binned in intervals of 4 points. The average score for the SCALE-UP class was (59.0 ±3.5)% for N = 42 students and the average score for the lecture class was (45.8 ±1.7)% for N = 162 students; these average values are indicated by the red arrows in the figure. Based on the standard error computed from the respective score distributions, the SCALE-UP class was observed to have performed statistically better than the lecture class.
Table 1. Summary of the data collected for our comparison between the SCALE-UP class and the lecture class. It should be noted that the only required element for all students in both sections of the Physics 1 class was the Final Exam in early 2018 (shown in red). Everything else was optional.

| Summative assessment          | common mid-term exam (week #10) |
|-------------------------------|---------------------------------|
| Diagnostic conceptual assessment | Force Concept Inventory (pre/post-tests) |
| Formative assessment          | 4 short conceptual/numerical quizzes |
| Feedback and comments         | mid-semester questionnaire       |
|                               | standard ETH course evaluation   |

Figure 3. Results of the mid-term exam for the SCALE-UP class (upper panel) and the lecture class (lower panel). The respective averages are shown by the red arrows in each case.

An item analysis of the mid-term exam was conducted to show the specific scores on each of the six exam questions, and these results are shown in Fig. 4. In each of the six questions, the SCALE-UP class had a higher score than the lecture class. It is also interesting to note that the margin is actually somewhat greater for the conceptual questions than for the numerical ones. For conceptual questions, the difference between the two classes varies from 1.7 to 2.6 points, whereas for the numerical questions, a difference of 0.7 to 1.4 points is observed. A more detailed analysis of the conceptual and numerical components
of the mid-term exam is presented in another GIREP paper (Schiltz et al. [10]) from this same conference.

The FCI [8,9] was given to both classes at the beginning and the end of the semester, and the results for these 30 conceptual questions are presented in Fig. 5. In the figure, the scores are binned in intervals of 2 points. The average values for each case are indicated by the red arrows in the figure, and these averages are summarized in the top panel of Fig. 6. By looking at the sample sizes for both cases, one obvious issue is an apparent disparity between the number of students who took the pre-test and the post-test. For the SCALE-UP class, this is not so severe ($N_{\text{pre}} = 44$ and $N_{\text{post}} = 37$). However, for the lecture class there is a very large discrepancy, where the pre/post populations differ by almost a factor of three ($N_{\text{pre}} = 153$ and $N_{\text{post}} = 57$).

Figure 4. Item analysis for the three conceptual questions (1C, 2C, 3C) and the three numerical questions (4N, 5N, 6N) from the mid-term exam. The SCALE-UP class is shown by the orange bars; the lecture class is shown by the green bars. Each question is worth 12 points.

One important consideration in the comparison between the SCALE-UP and lecture sections is the question of population equivalency. That is, one could ask the question: Does the SCALE-UP class consist of the self-selected “best” students? Without access to any student records, the only means by which we could address this question was using the FCI pre-test scores. We examined a sub-set of both class populations, restricting our comparison to only those students who had taken both the FCI pre-test and the mid-term exam [10]. This gave an average FCI pre-test score of $13.7 \pm 1.1$ for the SCALE-UP class ($N = 35$) and $12.2 \pm 0.6$ for the lecture class ($N = 92$). These results agree within the uncertainty and confirm that the SCALE-UP class was essentially equivalent to the lecture class.

For the FCI, the relevant indicator of student performance is the normalized gain $\langle g \rangle$, given by the relation $\langle g \rangle = (\text{post} - \text{pre})/(30 - \text{pre})$, which characterizes the fraction of missed points on the pre-test that were recovered in the post-test. In this case, we have calculated this quantity in two ways. First, if we simply find the average for all students who took the pre-test and the post-test separately, without restricting to matched pairs that have both pre/post results, we obtain gains of 0.48 and 0.37 for the
SCALE-UP and lecture classes, respectively. These results are shown in the bottom panel of Fig. 6 as the dark left-hand bars on each side.

Due to the disparity in participants between the pre/post-tests, we also opted to calculate the gains from only those students who had a matched pair for the pre-test and post-test. This is a more restricted data set, but with matched pre/post-tests we can compute the gain for each individual student and then take the average of that distribution. This leads to new results for the normalized gains which are 0.50 ±0.05 for the SCALE-UP class (N = 30) and 0.38 ±0.03 for the lecture class (N = 38). These results are shown in the bottom panel of Fig. 6 as the light right-hand bars on each side (with an error bar). In any case, it is clear that these new matched results do not differ substantially from the initial results (left-hand bars) where the average of the entire participant population was taken. In fact, both methods for computing the gain for each class agree within the uncertainty of the measurement, as seen by the error bars in the bottom plots. Either way, it is observed that the SCALE-UP class has a statistically higher normalized gain on the FCI than the lecture class.

![Figure 5. Results of the FCI for the SCALE-UP class (left panel) and the lecture class (right panel); the pre-test is on the top and the post-test is on the bottom. The averages are shown by the red arrows.](image)

We also note (see the bottom panel of Fig. 6) that the normalized gain for the SCALE-UP class is consistent with the range corresponding to an “interactive engagement” class, as determined in a paper by Hake [11] and indicated by the green band in the lower plots. While the gain for the lecture class does not quite reach that level, it is nevertheless higher than the “standard lecture” range indicated by the red band. This is probably due to the fact that there were some conceptual questions delivered to the lecture students, roughly 2-3 questions per class period, which gave them some chance for discussion of physics concepts.

Finally, we collected student comments in the middle of the semester to get a sense of how they were reacting to the novel pedagogical environment of the SCALE-UP class. Some representative comments from the students are presented below:

- I like that we work in groups. Trying to solve the given tasks together while exchanging helps a lot, you immediately learn from your mistakes. You can help others while others help you.
Figure 6. Summary of the FCI results for the SCALE-UP class (left panel) and the lecture class (right panel). The top portion of each panel shows the average scores for the pre-test and post-test. The bottom portion of each panel shows the normalized gain determined two ways (see text).

- I really do like doing and studying Physics the way we do, because through practical and daily-life examples I understand it better. It’s useful to have the possibility to come and learn by doing, rather than just hearing and copying. My only concern is the kind of asymmetry between our class and the lecture class: the exercises are different and what worries me a little is: "how is the exam going to be"?

- The class is great as one can really understand what we are doing but if you prepare using MasteringPhysics the class is not as difficult as the Moodle exercises online.

- I like the method of thinking we are taught in class. It helps to really understand the physics part of a problem and that it is not just the math behind it (as one could guess from the solutions to the Moodle problems). It might be a bit confusing that the Moodle problems aren’t always the same topics than what we have discussed in class. But I guess we will cover them later. Thank you all for making this opportunity possible!

- I love the atmosphere of the class, the different way of lecture! It is awesome to have such a motivated professor. Keep it up :)

- For me it is helpful if the professor gives me a small "relatable" knowledge. Through simple everyday experiences I can see physics working in my daily life. I build on that knowledge by doing simple steps that are logically connected. I follow first an intuition, how it could be, and afterwards I look mathematically to explain it. Before was only giving knowledge but not knowing how to use it.

- I think it’s a very interesting and good project, and I’m happy to take part in this SCALE-UP class. The only thing that may cause problems is when we do exercises in class if there is one person who already can solve the problem then he does all the work and the others don’t get time to think about the problem. In my opinion, it should be more that the ones who already know could maybe be a help but try to stay back a little bit and let the others think.

- I’m definitely learning physics but I’m not sure if I’m being prepared right for the exam as it seems that the Moodle problems are quite different from the ones we discuss in class. So I’m not sure if there will be the same focus in the exam as there is in the SCALE-UP class.
• It’s good to learn physics more practically and I think it will be more useful for everyday life than a lecture. It’s more interesting and I like to make the preparations for class by myself, so I can already summarize all the important things and formulas. For me, I can better learn the theory by myself than in a class of 200 students. I wish some of the exercises in SCALE-UP would be a little more like the exercises on Moodle, which we need to know for the exam, but I hope we’ll still keep the kind of questions we have at the moment.

While the majority of the comments were quite positive, there was some concern about possible differences between the two class sections, as reflected in the 2nd and 4th comments above (also see the last two comments). The activities in the SCALE-UP class tended to be more evenly balanced between conceptual and numerical problems; however, some of the students were worried that the lecture class seemed to emphasize numerical problems more heavily and that the Final Exam would focus entirely on this aspect. To address this concern, it was explained to the students that while the two class sections were following somewhat “different paths” through the semester, the coverage of the physics topics was the same in both cases and the ultimate goal of both sections was also the same. In the end, the SCALE-UP students would be well prepared (perhaps even better prepared?) for their Final Exam by having covered the conceptual aspects of the material in some detail in order to bolster their overall comprehension. This would provide a solid foundation in the physics concepts which would potentially lead to a greater facility in treating numerical problems.

5. Conclusions
In summary, we have conducted the first fully collaborative SCALE-UP class at ETH Zürich in an introductory physics course with 54 first-year undergraduate students. We found that the students were highly engaged during the class period, and they seemed to be well prepared for the classroom activities. The students were not at all hesitant to speak up during class – group members freely reported their results to the entire class and follow-up questions flowed naturally from the students during the class period.

A detailed comparative study was performed by collecting data from the SCALE-UP section and from a parallel lecture section. In order to ascertain the equivalency of these two student populations, FCI pre-test scores were used to verify that the students in both sections were essentially equivalent. In terms of assessments, FCI scores on pre/post-tests showed a higher gain for the SCALE-UP section as compared to the lecture section. Furthermore, results of an identical mid-term exam given to both sections also showed a higher average exam score for the SCALE-UP students. Feedback from the SCALE-UP students obtained via a mid-semester survey revealed a definite positive attitude towards the collaborative pedagogical approach. All of these results are very encouraging, with some caveats that are outlined below.

While the outcome of this study is favorable, it is worthwhile to note several constraints that possibly worked to hinder the positive benefits (to some extent). First of all, the actual contact time with the students for the SCALE-UP intervention was limited to only 3 hours on one class day per week (in fact, this is actually even less due to breaks, where $3 \times 50\text{ mins} = 150\text{ mins} = 2.5\text{ hrs}$). One extra hour for the recitation session, on a separate day, did not specifically involve collaborative activities since it was reserved for question/answer periods with the TA. By contrast, the contact time with introductory physics students in classes at George Washington University is typically 5 hours per week, which is significantly more. Second, the overall academic schedule for these ETH students is extremely intensive – they are typically in class for 30-32 hours per week. This makes it very difficult to expect them to work on homework problems outside of class, simply due to a lack of time. Third, since none of the assignments were required (or actually counted for a grade, except the mid-term exam, which essentially constituted a “bonus”), this contributes even more to a potential lack of incentive for practice outside of class. This issue also pertained to the level of participation in the assessments used for our comparative study – it is evident that the participation in our assessments (mainly the FCI and the mid-term exam) was highly variable due to the fact that these items were purely optional for the students.
6. Outlook

As a final commentary, one would observe that it seems that the academic culture tends to delay the “hard studying” until a period of time just preceding the Final Exam. For the SCALE-UP class to be even more successful, there is a need to provide time for the skills being developed in the classroom to evolve and “sink in” through a regular sequence of practice at home and feedback by the instructor. However, if this cycle is not followed, then the students are not really able to “close the loop” on the instructional benefits obtained from the collaborative setting. While the students’ performance in the SCALE-UP section was certainly impressive, and their level of satisfaction and sense of accomplishment was high, one could argue that even better results might be obtained if more time could be allocated to in-class activities and homework outside of class, as well as mandatory assignments that keep the students on track as the semester progresses. In addition, better data could be collected if participation in the assessments would be required for all students, even if those assessments do not count formally for their course grade.

Moving forward, it would be very interesting to try the SCALE-UP collaborative approach in another venue, with some of the improvements mentioned above. While the educational systems of various European countries can be rather different, the basic motivations of the students should be similar, in general, and we have found that the “novelty factor” for this type of collaborative group-learning environment was appealing to the students. Making another pilot study and collecting a robust set of assessment data would prove to be a big step in validating the effectiveness of this approach and confirming its universal applicability.

Acknowledgements

One of us (G.F.) would like to thank the Fulbright Specialist Program for support during the initial visit to Zürich in August/September 2016, as well as the Department of Physics and the Office of the Rector at ETH for support during the main period of this project (February–June 2017). The warm hospitality of the entire Physics Education Group, including my co-authors and other group members, was deeply appreciated. By graciously welcoming me as an integral part of this cohesive group, they all made my time in Zürich personally satisfying and professionally productive.

References

[1] McDermott L C 2001 Oersted Medal Lecture 2001: Physics Education Research – The Key to Student Learning Am. J. Phys. 69 1127–37. DOI: 10.1119/1.1389280
[2] Beichner R J et al 2007 The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) Project Research-Based Reform of University Physics ed Redish E F and Cooney P J (College Park, MD: AAPT)
[3] SCALE-UP web site is located at: http://www.ncsu.edu/PER/scaleup.html/
[4] TurningPoint web site is located at: http://www.turningtechnologies.com/
[5] Mazur E 1997 Peer Instruction: A User’s Manual (Upper Saddle River, NJ: Prentice-Hall)
[6] MasteringPhysics web site is located at: https://www.pearsonmylabandmastering.com/
[7] Giancoli D C 2009 Physics for Scientists and Engineers, 4th Edition (Upper Saddle River, NJ: Pearson Prentice-Hall)
[8] Hestenes D, Wells M and Swackhamer G 1992 Force Concept Inventory Phys. Teach. 30 141–58. DOI:10.1119/1.2343497
[9] Hestenes D and Halloun I 1995 Interpreting the Force Concept Inventory: A response to March 1995 critique by Huffman and Heller Phys. Teach. 33 502–6. DOI:10.1119/1.2344278
[10] Schiltz G, Feldman G and Vaterlaus A 2017 Active-learning settings and physics lectures: a performance analysis Proc. GIREP-ICPE-EPEC Conf. 2017 (Dublin) Journal of Physics: Conf. Series (in press)
[11] Hake R R 1998 Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses Am. J. Phys. 66 64–74. DOI: 10.1119/1.188809