

Teaching Physics Concepts without much mathematics: Ensuring physics is available to students of all backgrounds

Marcus. T. Wilson^{a*} orcid.org/0000-0001-6214-7727,

Sinduja Seshadri^b, orcid.org/0000-0002-3284-2910,

Lee V. Streeter^b orcid.org/0000-0002-3987-6945,

Jonathan. B. Scott^b orcid.org/0000-0001-8408-517

^aSchool of Science, University of Waikato, Hamilton, New Zealand;

^bSchool of Engineering, University of Waikato, Hamilton, New Zealand

*corresponding author: email marcus.wilson@waikato.ac.nz

Teaching physics concepts without much mathematics: Ensuring physics is available to students of all backgrounds

Recently, the University of Waikato introduced a first-year physics course requiring no specific physics or mathematics background, to support degrees in other subjects and allowing maximum diversity in student intake. This course uses the Investigative Science Learning Environment (ISLE) approach in which students develop their own experiments to tackle scientific questions. ISLE has proved successful with mathematically able classes; we have adapted it for a less capable cohort in a regional New Zealand context. We have assessed our approach for 2018 using: pre- and post-course concept tests; performance in exams; a student focus group; analysis of laboratory notebooks and marksheets; interviews with teaching staff; and student comments from the university's course evaluations. Concept tests showed a mean normalized gain of 0.61. Students did well in the exam. Many students were initially unfamiliar with practical scientific investigation and record keeping but laboratory notebooks showed desired learning outcomes were achieved by the end. Students and staff engaged well with this course. Students' final grades were not significantly different between genders and between under-represented minority and non-minority students. We cannot yet say whether this approach provides a sufficient platform for progress in engineering but students have learned fundamental practical physical concepts.

Keywords: Concept-index, concept inventory, physics, mathematics, teaching, Investigative Science Learning Environment

Introduction

Background

Underperformance of minority groups in engineering, physics and other science subjects has been identified as a barrier to achieving diversity in these areas (National Science Foundation, 2017; Eddy & Brownell, 2016). Minority groups include ethnic minorities, first-in-family-to-university and, in some subjects, gender minorities. Salehi,

Buckholder, Lepage, Pollock & Wieman (2019) have shown that the performance gap in physics is strongly linked to a gap in preparedness for university study. Mathematics is one such area in which incoming students are often underprepared (Rylands & Coady, 2009). However, merely making university courses more accessible to less prepared students by reducing reliance on assumed knowledge content does not in itself improve student learning and student beliefs about a subject (Elby, 2001).

Active learning has been demonstrated to be effective in delivering student learning outcomes in many subjects (Salehi et al., 2019; Ballen, Wieman, Salehi, Searle & Zamudio, 2017; Hake, 1998), and it is disproportionately effective for minority students (Haak, HilleRisLambers, Pitre & Freeman, 2011; Lorenzo, Crouch & Mazur, 2006). A major component of this effectiveness is likely due to increased student self-efficacy – i.e. how students see their own abilities (Ballen et al., 2017; Elby, 2001). However, active learning in itself is no guarantee of student success (Bernhard, 2017); one must also consider how students engage with the knowledge specific to the discipline (Case, 2019; Lorenzo et al., 2006).

In this work, we bring the assumption of limited background knowledge together with active learning. Specifically, we describe an attempt to create an introductory first-year university physics course that makes little use of assumed mathematical knowledge and uses active teaching methods. In particular, we ask the question: *“To what extent the combination of accessibility and active learning leads to success for first-year students who are poorly prepared for physics study?”*

The Physics in Context course

This ‘Physics in Context’ first-year undergraduate level course at the University of Waikato was introduced in 2018 to support students undertaking majors in other science areas, many of which enter with no explicit physics or mathematics background. There

are no subject-specific entry requirements above the basic numeracy and literacy standards needed for general entry to any New Zealand undergraduate degree. The course has therefore been put together so that mathematics is kept to a minimum. It does include rearranging formulae, sine waves, and the concept of vectors as a quantity with a direction. There is no calculus. While the course has been initially targeted at Bachelor of Science students taking majors in non-physics science subjects, it has also attracted would-be engineering students who are a substantial way off meeting the entry criteria for a Bachelor of Engineering (Honours) degree at the University of Waikato.

We have chosen the Investigative Science Learning Environment (ISLE) as our approach to active learning, as it has proven success in terms of students meeting learning outcomes across a variety of different, but mostly United States, institutions (Etkina & Van Heuvelen, 2007; Etkina, Murthy & Zou, 2006). ISLE is highly relevant to physics teaching because it emphasizes the approach to science taken in practice by science professionals including physicists, notably the cycle of observation, forming hypotheses, testing hypotheses and application of understanding, through experiments (Poklinek Čančula, Planinšič & Etkina, 2015; Etkina & Planinšič, 2014). In particular, it emphasizes the process of doing science, and applying scientific knowledge, as opposed to scientific knowledge itself. Broadly speaking, ISLE can be considered a problem-based learning environment, or inquiry-based science education, with small open-ended problems presented to students in laboratory classes. ISLE is ‘student-led’, meaning that explicit instructions for laboratory work are kept to a minimum; instead students formulate and carry out their own experiments to tackle a particular problem. ISLE also emphasizes physical concepts over mathematical equations, discouraging the “plugging numbers into equations” approach that leads to incorrect and unhelpful epistemological

thinking about the nature of physics and science (Wilson, 2014; Quale, 2011; Tumarino & Redish, 2007; Redish, Saul & Steinberg, 1998).

In the ‘Physics in Context’ course, students carry out a three-hour laboratory ISLE experiment in most weeks, along with three interactive lecture/tutorial sessions. Laboratory work is assessed, along with four assignments (one essay, one small-group oral presentation, and two based on short questions), two in-class tests and an end-of-semester exam. The laboratory class is where we intend most learning to take place and indeed to a large extent the lectures have been designed to support the learning undertaken in the laboratory sessions. In the laboratory sessions, students work in groups of three or four to tackle small experimental problems based on observation of phenomena, testing hypotheses and applying understanding, that is, engaging with knowledge. These problems are open-ended in the sense that detailed instructions are not given. Examples of observation and testing experiments are included in Appendix 1. Where possible, we have used ISLE “off-the-shelf” but we have needed to modify most experiments to be suitable for the lower mathematical ability of our student cohort. Students keep a laboratory notebook, which is assessed against a marking rubric. Content includes forces and motion, fluids, heat, electricity and magnetism, waves, optics, atoms and nuclei, elementary quantum physics concepts and elementary cosmology. Where possible, content is presented in a variety of contexts applicable across the breadth of science and engineering, for example seismic waves, Föhn winds, buoyancy of fish and chemical spectroscopy. We have intentionally used ethnically-diverse female names when using fictional characters in the laboratory manual, e.g. “Priya hypothesizes that....; design an experiment to test Priya’s hypothesis.” However, we have no way of directly evaluating what, if any, the impact of this is. The lecture and tutorial sessions did not follow a rigid ISLE structure. We chose to implement them as a

roughly equal mix of student-led problem-solving, and more traditional “chalk-and-talk”.

In 2018 there were 28 students enrolled initially, of which 23 finished the course. All those who dropped out did so early in the course, and also dropped out from other courses they were enrolled in. Of the remaining 23, five were female and eight were ‘under-represented minority’ students, for the purposes of this study defined as those who self-identified as Māori or were overseas students (i.e. not New Zealand citizens or Permanent Residents). While the definition of what constitutes a minority student can be debated, both these groups are known to be under-performing at the University of Waikato. Nine students out of the initial 28 enrolments stated informally that their ultimate intention was to study Engineering, others intended majoring in a wide range of subjects, in particular chemistry or computer science. The mathematics background of the cohort was limited. Only five students had obtained a level of performance in final year school mathematics that would have allowed them to enter first year engineering courses. All but one student had not obtained sufficient performance in final year school physics to enter engineering directly.

In this research we focus on the implementation of the ‘Physics in Context’ course in 2018 and on the ISLE approach in particular. ISLE has had proven success in large institutions amongst physics and other science major students. For example, ISLE has been shown in first-year university students to develop (a) high-level thinking and reasoning skills, (b) ability to design suitable experiments for a problem, (c) ability to transfer skills from one context to another, and (d) ability to critically analyse one’s own performance (Etkina, Van Heuvelen, White-Brahmia, Brookes, Gentile, Murthy, Rosengrant & Warren, 2006). Our context differs in that it is implemented with a group of students with little physics and mathematics background at a regional New Zealand

university. We have assessed student performance towards experimentally-based learning outcomes using the marking rubrics, and content knowledge through pre- and post-course concept tests. We have also analysed student feedback through a focus group and end-of-semester student evaluations, and considered feedback from the teaching staff.

Approach

This research was approved by the University of Waikato Faculty of Science and Engineering Human Ethics Committee.

We have considered several measures of student learning. First, we used pre- and post-course testing in 2018 to look at conceptual understanding of students before and after the course. Students were given a brief, ten question multiple choice (five responses, one correct) test before and after the course. The questions were developed by the authors in order to cover the full range of topics in the course, targeting what we thought would be “threshold concepts” (Meyer & Land, 2003). We acknowledge that these questions were not robustly tested beforehand, but were a first iteration. The questions are recorded in Appendix 2. The questions were all conceptual, involving no mathematics, along similar lines to the well-used Force Concept Inventory (Hestenes, Wells and Swackhamer, 1992) and the Signals and Systems Concept Inventory (Wage, Buck, Wright & Welch, 2005). These established inventories were not used because they covered between them less than half the content of the Physics in Context course. The test was implemented with a scratch-card so students had instant feedback as to whether their answer was correct. Multiple tries were allowed; ‘5’, ‘2’ and ‘1’ marks were given for answers that were right ‘first time’, ‘second time’ and third time’ respectively. The post-test assessed the same concepts as the pre-test, either by asking

an identical question (2 occurrences), asking the same question but in a different way (3 occurrences) or asking a different question about the same concept (5 occurrences).

Results were analysed in terms of actual marks and in terms of the fraction of the maximum possible gain, 'g', defined as: $g = [\text{Post-course score (\%)} - \text{pre-course score (\%)}] / [100\% - \text{pre-course score (\%)}]$ (Hake, 1998). Hence $g=1$ denotes that the student has achieved the maximum possible gain and $g=0$ that no gain has occurred. It is possible for g to be negative. The 'g' is useful since it allows more direct comparison between students with different incoming knowledge base than a simple percentage mark gain. We chose to give the test to all students, not a subset of students, since enrolments were small; thus we note we have no control group for comparison.

Secondly, we have assessed the 2018 students' acquisition of laboratory skills, through the marking rubrics. Separate rubrics, as presented by Etkina et al., (2006) and Etkina & Van Heuvelen (2007), were used for general laboratory skills, observation experiments, testing experiments and application experiments. More than one rubric was relevant to each laboratory session. In addition, we developed our own general laboratory skills rubric along similar lines to the others, which we include in Appendix 3. In particular, we have focused on the laboratory skills and testing experiment rubrics because these are used in every laboratory experiment. The laboratory skills rubric contained the criteria:

- Ability to keep a good record of what was done;
- Ability to demonstrate thinking behind the chosen methods;
- Ability to make accurate observations and measurements;
- Ability to present results clearly;
- Ability to make sense of results.

The testing experiment rubric contained the criteria:

- Ability to identify the hypothesis;
- Ability to distinguish between hypothesis and prediction;
- Ability to make a reasonable prediction;
- Ability to assess an outcome against a prediction;
- Ability to make a reasonable judgment about a hypothesis.

Each skill was determined to be ‘missing’, ‘inadequate’, ‘needing improvement’ or ‘adequate’; halfway points between these were allowed. We identified how many students were meeting the ‘adequate’ criteria, or halfway between ‘needing improvement’ and ‘adequate’, as the semester progressed.

Student experiences with this course were collected via online end-of-semester evaluations and a focus group. The evaluations asked for responses to the following statements, common across all University of Waikato courses:

1. My experience of the course matched the information on the course outline;
2. The learning opportunities helped me to achieve the learning outcomes;
3. The assessment tasks supported my learning;
4. The content helped me to extend my knowledge;
5. The course design helped me to examine different theories, concepts and practices;
6. The course was well organised;
7. Overall the course provided me a good learning experience.

Students responded to each with one of five options: ‘Strongly Agree’, ‘Agree’, ‘Disagree’, ‘Strongly Disagree’, or ‘Not able to judge’. Additionally, students were

asked for free responses to the questions: ‘What aspects of the course helped you learn?’ and ‘What improvements could be made to this course?’.

The focus group was run by a staff member not connected with the teaching of the course, after teaching had finished, but before the final examination. Students were asked to comment on the course content and structure, how the course compared with other courses, their overall experience with it and how it could be improved.

Teaching staff have also discussed their experiences. Teaching was shared between two lecturers, and one graduate student assisted in the laboratory classes.

Results and Discussion

We start by presenting the quantitative measures of student performance, then move to the qualitative indicators of student and staff experiences with the course.

Pre- and post-course testing

Seventeen students completed both the pre-course test and the post-course test. Figure 1 shows results. In Fig. 1a we present the post-course mark against the pre-course mark for these seventeen students. All students except one improved in the post-course test; the other student achieved an identical mark as in the pre-course test. Figure 1b shows the same data in terms of the normalised gain, ‘g’. The mean normalised gain for the students was 0.61, indicated by the thick dotted line. This is towards the high end of the “medium” range of ‘g’ identified by Hake (1998), which used the Force Concept Inventory to assess mechanics. The actual distribution of marks for the pre- and post-course tests are shown by histograms in Fig. 1c and 1d respectively. The two distributions are significantly different ($p < 0.001$ with a paired t-test) with data consistent with normal distributions ($p = 0.199$ and 0.194 with an Anderson-Darling test

for pre- and post-tests respectively). It is important to note that we have not demonstrated that improvements are a result of ISLE. Possibly the post-test was easier. Gains might also be expected with traditional teaching approaches.

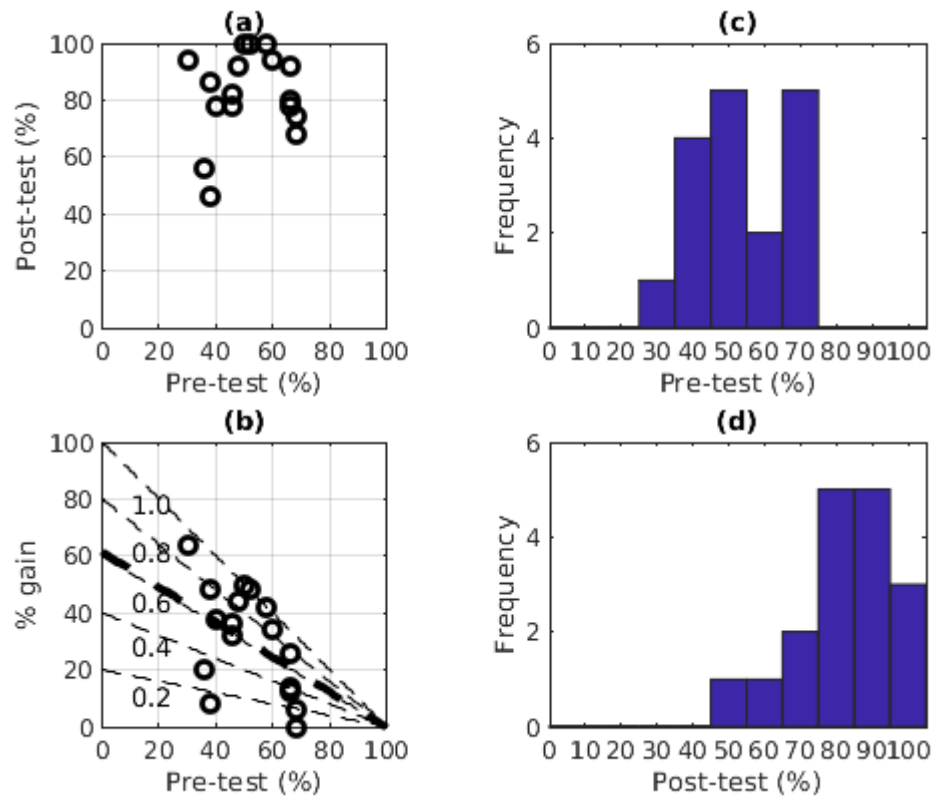


Figure 1: Results of the pre- and post-tests. (a) The post-test score of each student against their pre-test score. (b) The percentage gain of each student against their pre-test score. The thin dotted lines denote normalized gains of 0.2, 0.4, 0.6, 0.8 and 1.0. The thick dotted line denotes the mean normalized gain of 0.61. (c) A histogram of the pre-course test results. (d) A histogram of the post-course test results.

Table 1 shows how the improvement was broken down against topic, which is ordered in terms of chronological order of teaching the topic. Greatest improvement was seen in the topics taught towards the end of the course; in contrast there was little improvement in student marks in the two mechanics questions. This contrasts with some

other implementations of ISLE assessed with the Force Concept Inventory (Etkina & Van Heuvelen, 2007), which show larger gains. In Physics in Context, mechanics was taught only in the first two weeks of the course, so this difference might be due to the most recent topics being better recalled at the time of the post-test. However, we note that the end-of-semester exam saw an opposite trend; those questions on topics taught earlier in the course were answered better. Interestingly, the Optics & Quantum Physics and Nuclear Physics & The Universe questions gave ‘g’ scores larger than the mean exam mark, indicating disparity between these measures. However, we note that the questions were arranged on the exam paper in the order of Table 1 and therefore the low exam marks for these two questions might be a result of them being tackled at the end of the exam when students were under time pressure. Two post-test questions were identical to pre-test questions, one in Mechanics and one in Nuclear Physics & The Universe. While this might explain the high ‘g’ for the latter section it also implies students did not benefit in the same way in the former section.

Table 1: Gain in score for each topic averaged across students and questions on the topic, compared with the mean exam mark for each topic.

Topic	Question numbers in pre-test; and post-test	Mean 'g'	Mean exam mark (%)
Mechanics	3 & 6; 1 & 7	0.22	71
Thermodynamics	2 & 4; 3 & 10	0.55	65
Electricity & Magnetism	10; 6	0.60	68
Vibration & Waves	7; 9	0.40	70
Optics & Quantum Physics	5 & 8; 4 & 5	0.87	61
Nuclear Physics & The Universe	1 & 9; 2 & 8	0.85	41

Acquiring laboratory skills

The skills acquired by students during the laboratory sessions are presented against chronological laboratory number in Figure 2. There was no ‘rotation’ system of laboratories; in a given week all students were doing the same laboratory experiment. Each point on the plot shows the fraction of students who had demonstrated, or nearly demonstrated, each of the specific abilities on the marking rubrics (that is, had been assessed as ‘adequate’ or between ‘needs improvement’ and ‘adequate’ through the rubric.) Figure 2a shows how students acquired general laboratory skills; the five lines corresponding to the five abilities assessed by the rubric. Although there is considerable variety, one can see a general increase in demonstration of these skills as time progresses, particularly for the ability ‘to make sense of results’. Figure 2b shows how students tackled ‘testing’ experiments. Again, there is a general increase in skills as time goes on. Graphs are not shown for ‘observation’ and ‘application’ experiments since some weeks did not feature these experiments. As with the pre- and post-testing, we cannot say that these gains are solely a result of ISLE, but we remark that ISLE has been designed to achieve such gains.

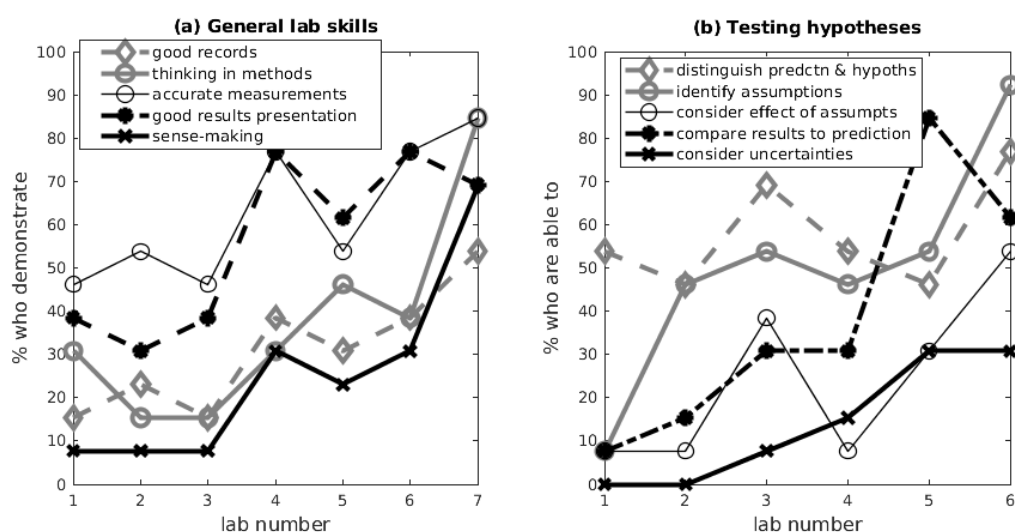


Figure 2: Acquiring of specific laboratory skills. (a) Plots of the percentage of students who demonstrated, or almost demonstrated, five specific general laboratory skills, against the lab number. The first lab students did is marked as '1', the second '2', and so forth, so that the x-axis represents increasing time into the semester. (b) Plots of the percentage of students who demonstrated, or almost demonstrated, five specific skills related to testing a hypothesis, against the lab number.

Figure 3 shows the distribution of laboratory marks against laboratory number. The plot shows lower quartile, median and upper quartile marks. Students have improved as semester has progressed, particularly over the first three laboratory sessions.

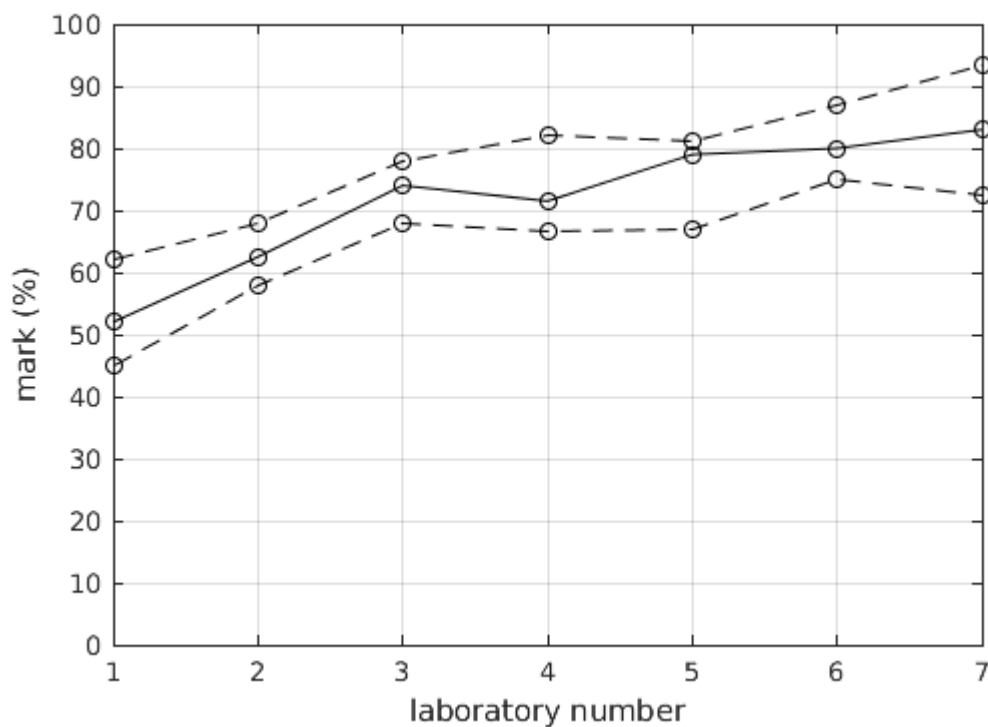


Figure 3. Laboratory marks (%) as a function of laboratory number. The solid lines show the median results; the dashed lines show the lower and upper quartiles.

Overall marks

Twenty out of the 23 students who finished the course passed the course. Figure 4a shows the breakdown in overall marks for the Physics in Context course, split between male and female students. There were no significant differences between males and females ($p=0.38$, with an unpaired two-tail t-test.) There were also no significant differences between the groups in any of the laboratory marks as semester progressed. Figure 4b shows a similar plot comparing minority and non-minority students. Again, there are no significant differences between the overall marks for the two groups ($p=0.72$). We note that there were no observed differences in mathematics or physics background between these student groups. We can conclude that in our implementation ISLE has not disadvantaged minority students, but not that it has disproportionately benefited them as found by Haak et al. (2011) and Lorenzo et al. (2006). While a better measure of the effect on minority students might be the comparison between post-course and pre-course tests, we note that just two minority students completed *both* the pre-course test and the post-course test, meaning we cannot reasonably consider these data as representative.

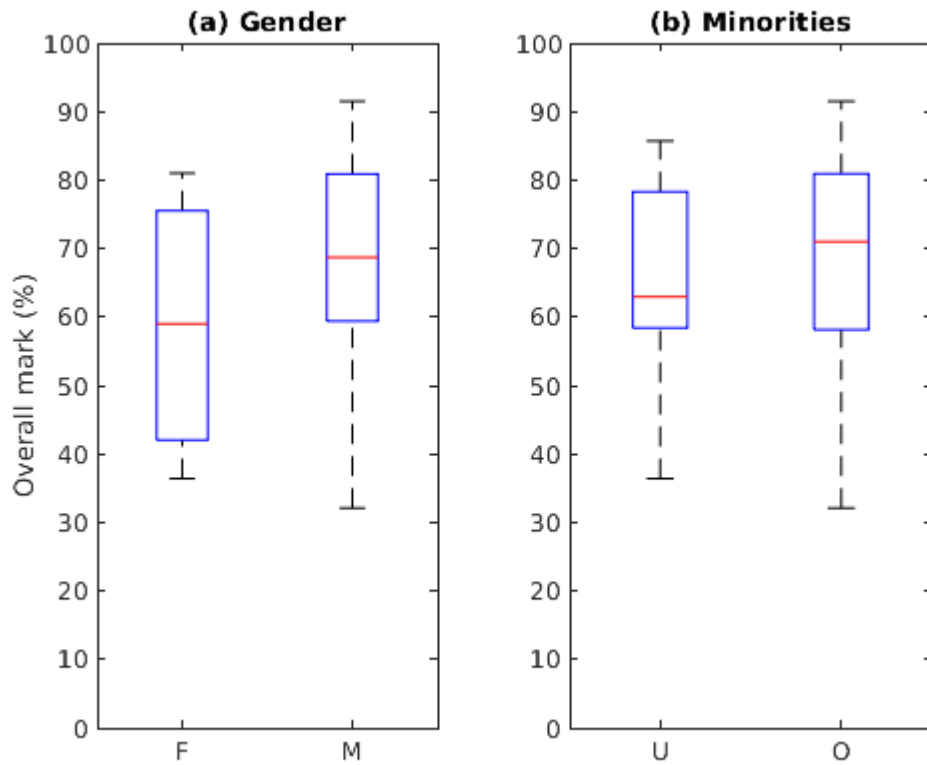


Figure 4. The distribution of final marks (as percentage) in 2018 for (a) female (F) and male (M) students and (b) under-represented minority (U) and other students (O). The central line denotes the median; the boxes denote the interquartile range; the whiskers the extreme points in the datasets. There are no significant differences between the groups.

Focus group

We now report key messages from students arising from the focus group. There were eight students present of which four were female. We did not collect data on minority background of the students. We acknowledge that these are specific experiences of a few students and not necessarily representative more broadly. Overall, students thought they had achieved a lot intellectually. One student commented “I did not expect to get so much...”, another:

“I learnt that I thought I understood, but then I learnt that I didn’t, and now I think I do.”

In particular, students recognized that a strong focus of the course was laboratory work and that they achieved a lot in this regard. However, in contrast to their experience at school, they felt that the course placed emphasis on making sense of what was done in a particular context. One reported

“[This course is] about describing everyday things in physics language,”

which echoes the intention of the course, being “Physics in Context” and recognizes the role of language in physics teaching (Brookes & Etkina, 2015).

The keeping of a laboratory notebook was a particular point of difficulty. At the start of the course, many focus group students reported having little idea as to what should and should not be recorded in the book. Students felt they needed more explicit teaching in this regard. Also, formulating hypotheses and making predictions based on them was challenging, as this is not usually tackled at secondary school.

The students wanted more time to be spent in the laboratory; one three-hour slot was seen as insufficient for each experiment. Students asked also to have the chance to use more advanced equipment. Students also reported achieving three learning outcomes that had not been explicitly stated: 1. How to record scientific data, 2. How to break down a problem into smaller parts, and 3. How to form a hypothesis and test it. Students expressed surprise at the overall lack of mathematics in the course, although they made no requests for it to be increased.

End of semester evaluations

There were fifteen completed on-line evaluation forms from the 2018 cohort. Of the seven questions asked, six had “agree” as the modal response. The other question, “The course design helped me to examine different theories, concepts and practices” had a modal response “strongly agree”, perhaps echoing the focus group comments that the course was not primarily about content knowledge. However, we recognise research suggesting that students might be unable to distinguish between different questions in an evaluation (Wieman, 2015) and indeed that little correlation exists between student evaluations and actual learning (Uttl, White & Gonzalez, 2017).

When asked “What aspects of the course helped you learn?” there were several comments around the ‘real-world’ context of what was taught:

“I enjoyed the real world applications of what we were learning”

“...what was being learned was very interesting and fun to participate [in]”

“The lab sessions worked well to contextualize the content and build an understanding of real life situations.”

When asked “What improvements could be made to this course?” ten responses strongly focused on the laboratory component, and six of these were asking for increased clarity on what should be written in notebooks. The response:

“...in the beginning [writing in the notebooks] was a hindrance to learning”

demonstrates the student view that the time taken *writing about* experiments was

detracting from the time taken *doing* experiments.

There was only one comment regarding the mathematics in the course:

“A lot of time was dedicated to teaching what was already easy maths”

Teacher experiences

We next briefly outline the self-reported experiences of the teaching staff, but note these may be specific to the individuals concerned. The various teaching staff reported that students were generally well engaged with the class, especially during the laboratory activities. One measure of this, but not explicitly assessed, was the low observed usage of mobile phones for non-laboratory-related activity. Peter, Khoo, Cowie, Scott & Round (2017) report that in Engineering classes even the best performing students spend about 8% of time engaged with an external activity on a mobile device during a laboratory class; but there was little evidence of this. A laboratory assistant commented:

“I didn't notice anyone using [their] mobile phones for personal use like text messages etc. during the lab, except for [a] work purpose [such as a] stopwatch;”

and this experience was backed-up by an external observer.

The experience of seeing students learn through thinking and trying out experimental problems for themselves, and critically appraising the outcome of their own experiments, was rewarding for the teachers, since it contrasted with experiences in many other courses they had taught.

There were several challenges faced. First, the varied background of students meant that for most sections of the course there were students who were familiar with the material but also students who were not. Also, some sections of the course resonated with some groups of students, but others less well. Mathematics ability was highly variable among students, as expected given there were no pre-requisites. A few students were familiar with elementary calculus, while others routinely made mistakes with rearranging simple equations (e.g. “rearrange $c = f \lambda$ to find λ ” was not always done correctly).

Considerable effort and perseverance was required by teachers early in the course, as it was clear that many students had little or no practical experience with science. For some, the concept of doing an experiment appeared new to them, and certainly the distinction between observation, testing and application experiments (Etkina and Planinšič, 2014) was not understood. Formulating and testing a hypothesis was particularly problematic. The statement “If hypothesis H be correct, when we do experiment X we will observe Y” was challenging for many students, especially early on in the course. There was some improvement with time, as evidenced by Figure 2b which show an improvement in students’ abilities to distinguish hypothesis from prediction (gray-dashed line) and interpret their results in light of predictions (black-dashed line). We found that students performed better when testing experiments could be put into clear *quantitative* terms. For example, students performed well when asked to test whether a light-bulb obeyed Ohm’s Law; they were able to see that a graph of voltage against current was not a straight line and thus conclude that Ohm’s Law was not obeyed. But *qualitative* problems were much more challenging; e.g. testing whether the size of the electrostatic force between two objects changed with distance between the two objects needed considerable teacher input. This echoes the much-quoted

findings from Mazur (1997) that students could competently use mathematical calculations to find currents and voltages in an electrical circuit, but could not correctly answer simple conceptual questions about the same circuit.

The question of what should be written in a laboratory notebook was always apparent and a constant challenge. Many students had considerable difficulty in translating their thoughts and understanding into words: “*Writing [lab books] is too much work*” summed up one student’s feelings. Brookes and Etkina (2015) highlight that language can influence how students conceptualize physics topics (in their example, heat) and recommend introducing technical language only when “*the how and the why of the phenomenon*” is familiar to them. Our experience is that more time and explicit teaching needs to be devoted to how science understanding can be put into words.

Teacher effort was also required to get students to talk to others in different groups. In ISLE students are asked to visit other groups to see what they had done, to share thoughts and experiences, and to learn from each other. There was often considerable reluctance to do this. This might be a cultural reluctance based around students not wishing to appear inferior to the peers (‘whakamā’ in Māori) although the reluctance did not seem to be related to ethnic background.

Conclusions

In this work we asked to what extent the combination of accessible entry requirements and active teaching leads to student learning in introductory physics. We have shown that students with low mathematical ability have been able to meet learning outcomes and acquire conceptual understanding in physics. While we cannot robustly say this was the result of our pedagogy, results are consistent with reported outcomes of ISLE in

different contexts (Etkina et al., 2006). In particular, students have acquired laboratory skills that they did not have at the start of the course. No differences in performance were observed between genders, or between under-represented minority and non-minority students, although this may be because numbers of students were low. Nevertheless, there is reason to believe that ISLE, as a student-led course, would be effective in retention of diverse groups of engineering students (Haak et al., 2011; Lorenzo et al., 2006). How effective this approach is for would-be engineering students in the long term has not yet been established, but we note in closing that of the nine students who initially indicated a goal of studying engineering, four are currently (as of early 2020) enrolled in a Bachelor of Engineering (Honours) degree at The University of Waikato.

Appendix 1

Example ISLE Experiment

The following is an example of the text provided to students for observation and testing experiments in a three-hour ISLE laboratory. This experiment has been modified by the first author from the “off-the-shelf” ISLE experiment on fluid heating and cooling available at <https://sites.google.com/site/scientificabilities/> in order to reduce the mathematical content.

OBSERVATION EXPERIMENT: TEMPERATURE CHANGES WHEN OBJECTS OF DIFFERENT TEMPERATURES COME TOGETHER.

Available equipment: kettle, containers, water, thermometers, metal blocks

Use the kettle to heat up a volume of water (you don't need to take it all the way to boiling). Pour the hot water carefully into a container, but don't fill the container all the

way up. Measure its temperature. Now, carefully place into the water one of the metal blocks. Measure the temperature change in the water. Repeat with other blocks. What influences how much the temperature of the water drops by when the block is added? Explain what is happening in terms of conduction and specific heat capacity.

Questions:

- a) How would you ensure each test is comparable with the others?
- b) What heat transfer processes occur when the block is added?
- c) When the block and water have reached equilibrium, do they have the same *internal energy*? The same *temperature*?
- d) How does the heat flow influence the temperature?
- e) Which factors are important? Explain why or why not:
 - i) The material the block is made from
 - ii) The mass of the block
 - iii) The volume of the water

TESTING EXPERIMENT: COOLING OF A CUP OF TEA

The goal of this experiment is to test a hypothesis associated with the cooling of fluids.

Available equipment: electric kettle, mugs, cold water from a fridge, measuring jugs, thermometers

Erin and Priya are discussing the problem of fitting a cup of tea into a short fifteen

minute tea-break at work. They find that after they have made the tea that the tea is still too hot to drink easily by the end of the tea break. They are discussing how to cool it as quickly as possible. Both like their tea with milk added. Should they add the milk immediately after making the tea, or let it cool first and then add the milk, or does it not matter? (Neither are particularly fussed as to what their tea tastes like!)

Erin hypothesizes that it is best to wait as long as possible to add the milk to the hot tea. She says that the hotter the fluid, the more quickly it will cool, so by letting it cool without milk to start, then adding cold milk just before you want to drink it, will get it as cool as possible. Priya is unconvinced – she argues that there is a certain amount of energy to start with, and a certain amount of time before the tea break is over, and so the tea will end up at a certain temperature at the end regardless of what order things are done in.

Carry out an experiment to test Erin's hypothesis that it is better to add the milk at the end.

Think about the following things:

- a) What will you measure? How will you do it? Write it down before doing the experiment. Be explicit.
- b) How many experiments will you need to do? How can you ensure that all experiments are equal ('fair')?
- c) How will you know if the evidence supports the hypothesis? Or not?

Appendix 2

Pre-course test

The pre-course test is reproduced below. Students answered all ten multiple-choice questions. The correct answers are indicated.

1. An atomic nucleus contains positively-charged protons and uncharged neutrons.

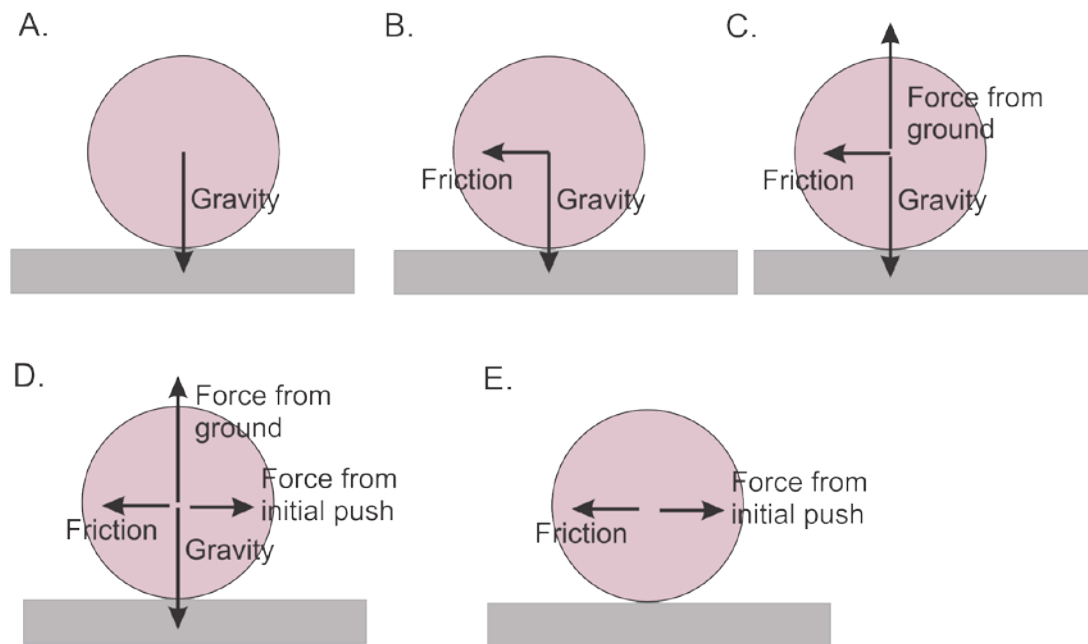
Positive charges repel, and neutral ones do nothing, so the nucleus should fly apart. So why does the nucleus stay together instead of flying apart? [Answer: B]

- A. Gravity
- B. There is another force, an attractive one, that is stronger than the repulsive force, but only happens in the nucleus.
- C. The positive-positive repulsion doesn't happen at subatomic distances
- D. The surrounding electrons shield the positive charges from each other
- E. Quantum coherence

2. If warm air rises, why is it colder at the top of a mountain than the bottom? [Answer: B]

- A. There is normally more cloud cover at the top of the mountain so it receives less sunlight
- B. Rising air expands because there is less pressure and this expansion takes heat energy from the air
- C. There is less atmosphere at high altitude to blanket the ground and keep it warm
- D. Cosmic rays are absorbed by the atmosphere at high altitude taking energy away from the air
- E. Mountain tops are exposed so there is more cooling by the wind than in a valley

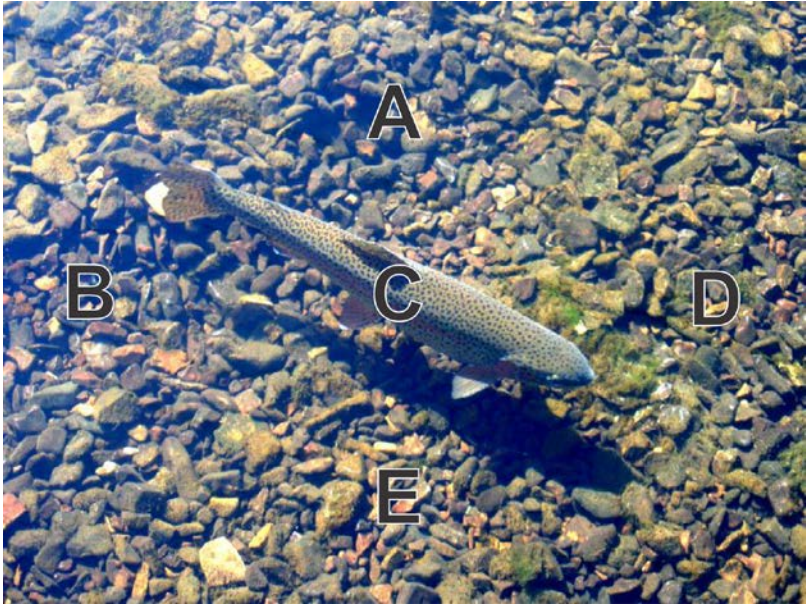
3. A bowling ball rolls along a carpet, from left to right. Which of the five diagrams (A-E) shows the forces acting on the bowling ball. [Answer: C]



4. I go swimming in an open air, heated swimming pool. The air temperature outside is 20 Celsius, but the swimming pool is heated to 28 Celsius. Yet I feel cold when getting into the water. Why? [Answer: D]

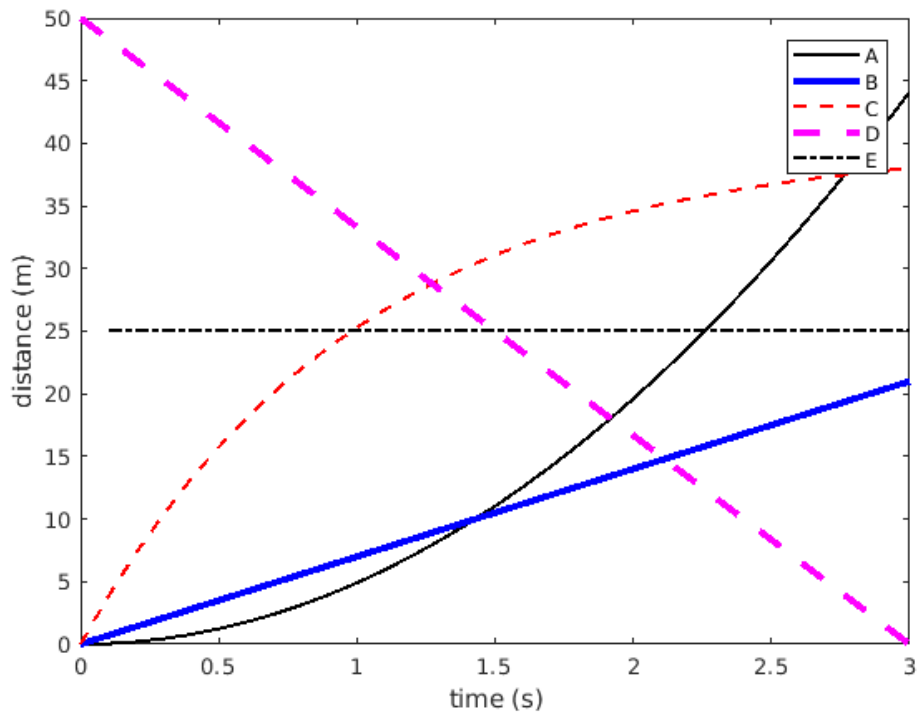
- A. Because water has a much higher specific heat capacity than air
- B. I only perceive the water to be cold - it's a psychological response caused by the skin becoming wet
- C. 28 Celsius water is actually colder than 20 Celsius air - it's because the temperatures of liquids and gases are defined in different ways
- D. Because water conducts heat from my body very easily
- E. The immersion of a partially-buoyant body in a fluid medium implies adiabatic expansion and thus radiative cooling of the medium

5. Gustav is hunting fish in the Upper Waikato with his bow and arrow. He spies a trout submerged in the water. Where does he aim his arrow? [Answer: E]

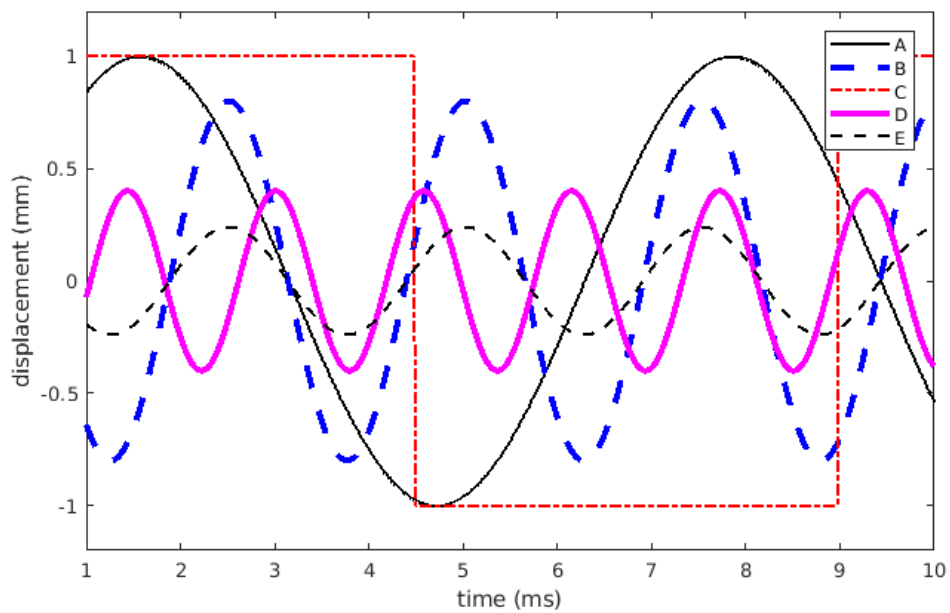


6. The five graphs (A-E) show the distance travelled by different objects as a function of time. Which one of them represents the motion of a large stone falling down a well?

[Answer: A]



7. The following figures show five waveforms, plotted as displacement (y-axis) against time (x-axis). Which has the largest frequency? [Answer: D]



8. Rather than a whole spectrum of colours, a hydrogen vapour lamp produces just four discrete visible spectral lines. Why? [Answer: B]

- A. Because there are four different isotopes of hydrogen
- B. Because the electrons in a hydrogen atom can only move between discrete energy levels
- C. Because the full spectrum of colours for hydrogen actually appears in infra-red which we can't see with our eyes.
- D. Because all the other colours of the spectrum are absorbed by the hydrogen vapour - these four lines are the only ones that get through.
- E. Because a hydrogen molecule has four modes of movement - three translational modes (one for each of the three dimensions) and one vibrational mode.

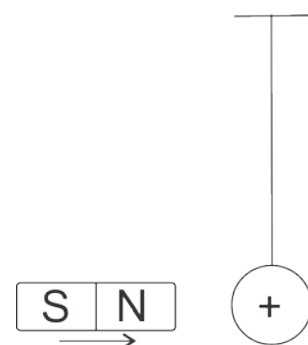
9. Billions of dollars (in 2018 equivalent) have been spent on research into nuclear fusion for each of the last 50 years, but without major success. Why does this funding continue? [Answer: D]

- A. Because it allows the U.S., Russia and other countries to make more destructive weapons
- B. Because scientists want to know more about the structure of the atomic nucleus
- C. Because it would tell us more about how stars (such as the sun) work
- D. Because if fusion can be achieved on a large scale we will have an almost unlimited supply of clean energy
- E. Because it is an exciting topic in physics

10. An acrylic ball is hung from a thread. It is given a positive charge. A bar magnet is brought up to the ball, with the north pole towards the ball, as in the diagram. What happens to the ball? [While the intention of the question was to test

whether students distinguished the magnetostatic force from the electrostatic force (hence C, no effect) we acknowledge with

hindsight that charge separation would occur in the magnet and hence an electrostatic attraction would occur.]

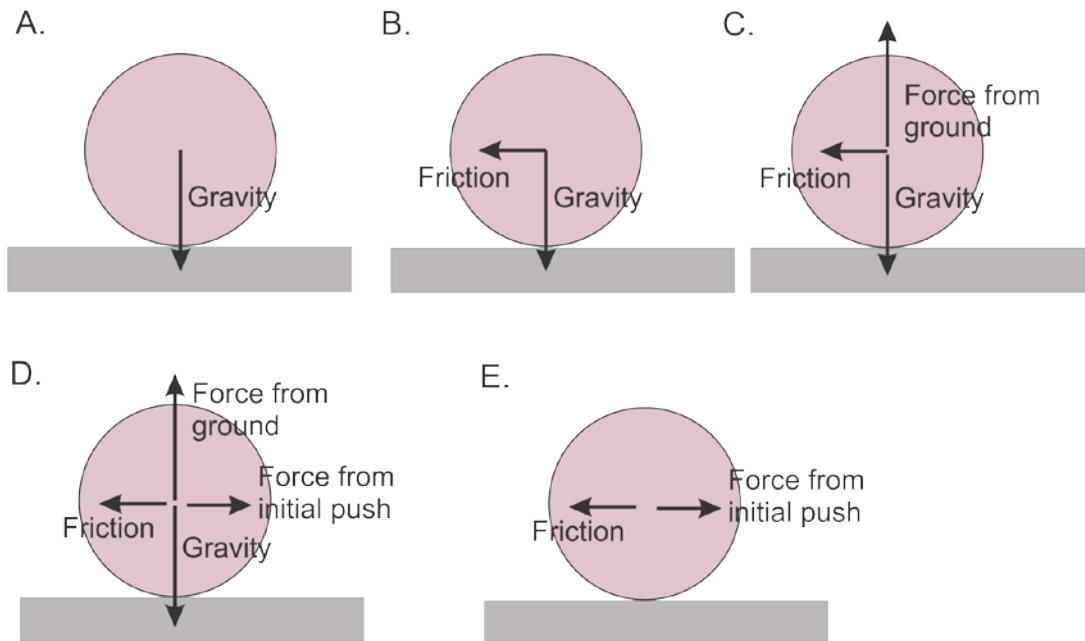


- A. It is attracted towards the magnet.
- B. It is repelled from the magnet
- C. Nothing
- D. The ball will start to swing to-and-fro
- E. The ball heats up

Post-course test

The post-course test is reproduced below. Students answered all ten multiple-choice questions. The correct answers are indicated.

1. A bowling ball rolls along a carpet, from left to right. Which of the five diagrams (A-E) shows the forces acting on the bowling ball. [Answer: C]



2. Billions of dollars (in 2018 equivalent) have been spent on research into nuclear fusion for each of the last 50 years, but without major success. Why does this funding continue? [Answer: B]

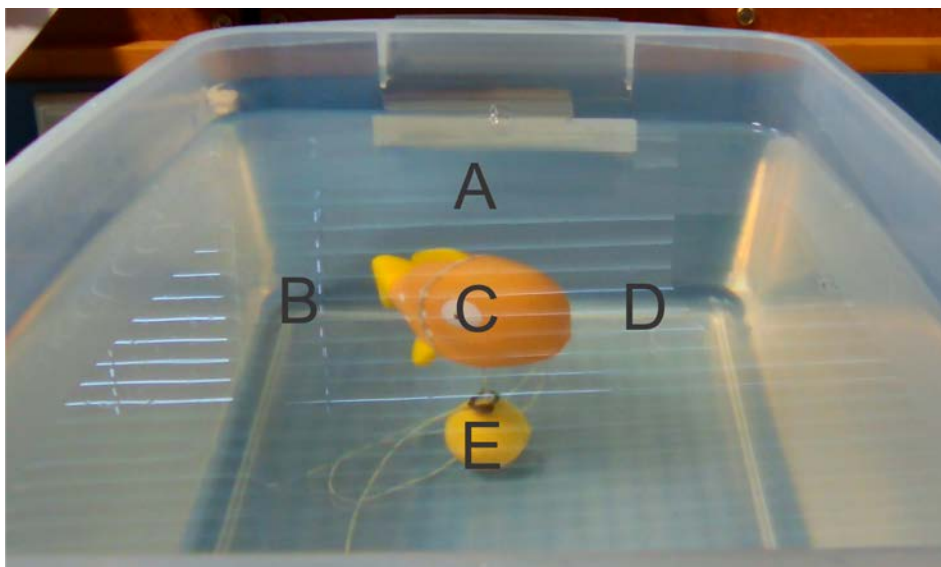
- A. Because it is an exciting topic in physics

- B. Because if fusion can be achieved on a large scale we will have an almost unlimited supply of clean energy
- C. Because it allows the U.S., Russia and other countries to make more destructive weapons
- D. Because it would tell us more about how stars (such as the sun) work
- E. Because scientists want to know more about the structure of the atomic nucleus

3. I heat my house to a constant 21°C . Yet when I walk with bare feet, the tiles on the kitchen floor feel colder than the lounge carpet. Why? [Answer: A]

- A. Because the tiles have a higher thermal conductivity than the carpet.
- B. The tiles and the carpet are actually at different temperatures – the tiles will be colder
- C. Because tiles have a higher density than carpet.
- D. There is greater friction between my feet and the carpet than between my feet and the tiles.
- E. It's only a perceptive illusion – the smoothness of the tiles make them feel colder than the rougher carpet

4. Erin is spearing toy fish in the laboratory. She has the following view. Where does she aim her rod to hit the fish? [Answer: E]

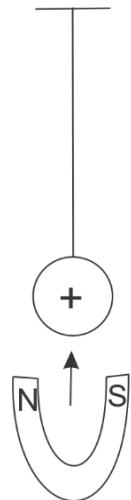


5. The light from the sun shines through a prism and it is split into a spectrum of colours. When light from a hydrogen vapour lamp shines through a prism only four very narrow bands of colour are seen. Why is there this difference? [Answer: D]

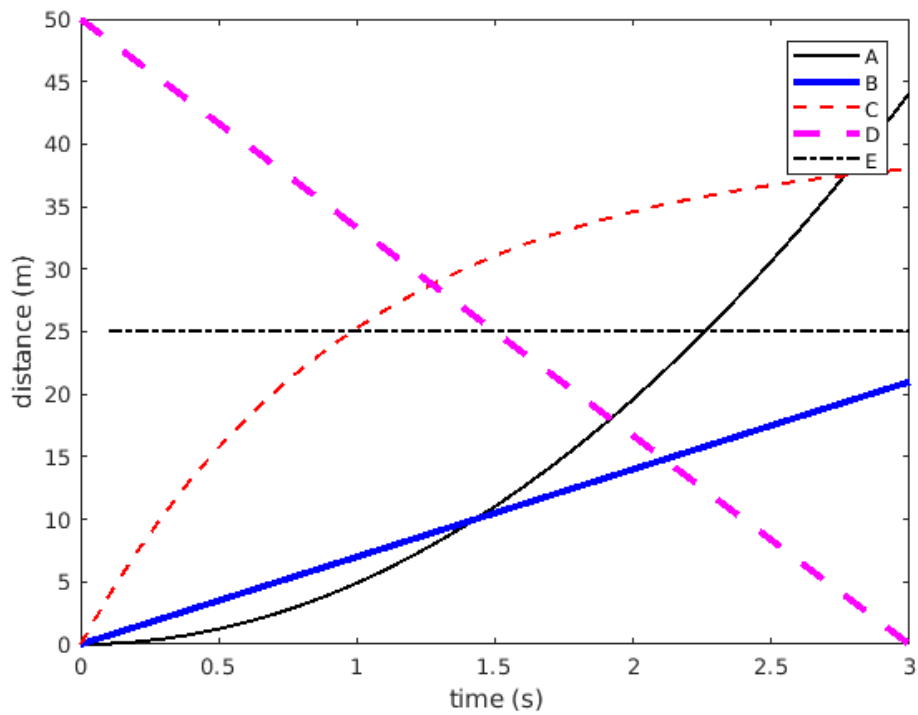
- A. Because there are four different isotopes of hydrogen in the vapour lamp, whereas there are thousands of different isotopes in the sun
- B. Because the hydrogen vapour Doppler-shifts most of the spectrum into the infra-red, and only a few lines remain
- C. Because a hydrogen molecule has four modes of movement - three translational modes (one for each of the three dimensions) and one vibrational mode
- D. Because the electrons in a hydrogen atom can only move between discrete energy levels, emitting very specific frequencies of light
- E. Because all the other colours of the spectrum are absorbed by the hydrogen vapour - these four lines are the only ones that get through.

6. An acrylic ball is hung from a thread. It is given a positive charge. A horse-shoe magnet is brought close to the ball, as shown in the diagram. What happens to the ball? [Answer: D]

- A. It starts to swing to-and-fro
- B. It moves left
- C. It moves right
- D. Nothing
- E. It heats up



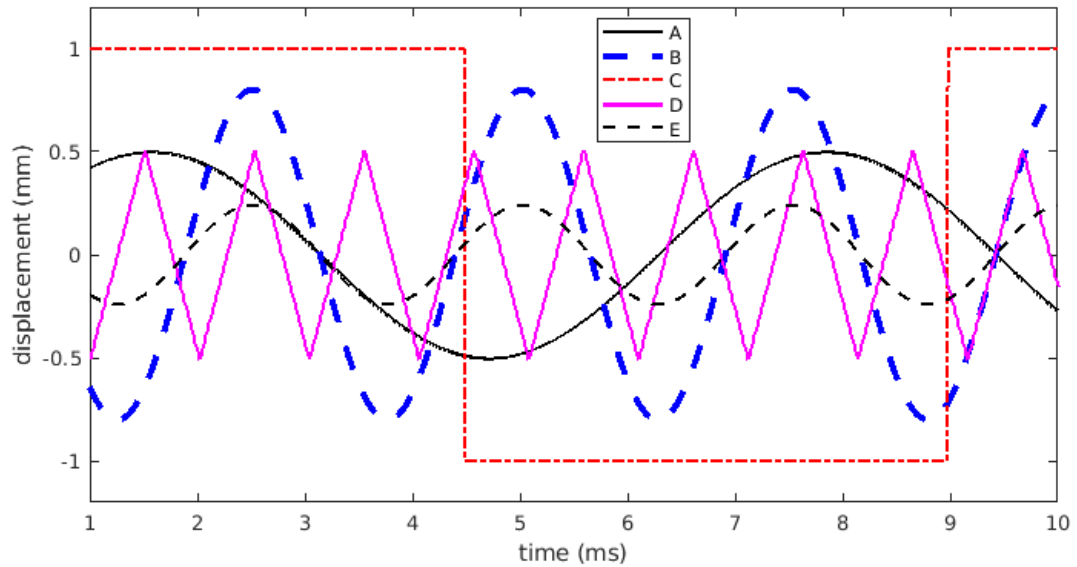
7. The five graphs (A-E) show the distance travelled by different objects as a function of time. Which one of them represents the motion of a feather falling in a vacuum chamber? [Answer: A]



8. What holds the nucleus of an atom together? [Answer: B]

- A. Gamma radiation
- B. The strong nuclear force
- C. Quantum coherence
- D. The electrons orbiting the nucleus
- E. Gravity

9. The following figures show five waveforms, plotted as displacement (y-axis) against time (x-axis). Which has the largest frequency? [Answer: D]



10. Why does the nozzle of an aerosol can become cold when you spray it? [Answer: B]

- A. The liquid inside is cold so the nozzle will cool as the liquid passes through.
- B. The liquid inside expands quickly into a gas as it passes through the nozzle, and this expansion takes heat away from the nozzle.
- C. Because the nozzle gets damp – and damp things are colder than dry things.
- D. When you press the nozzle down to spray you have to push hard, taking heat energy away from the nozzle.
- E. An endothermic chemical reaction between the liquid inside and the air outside takes place when the aerosol is sprayed.

Appendix 3

Laboratory Skills Rubric

Table 2 show the rubric used to assess students' laboratory skills. Rubrics for testing and application experiments are available at

<https://sites.google.com/site/scientificabilities/>

Table 2: The marking rubric used to assess general laboratory skills.

General skills: Ability to collect and analyse experimental data				
Scientific Ability	Missing (0)	Inadequate (1)	Needs some improvement (2)	Adequate (3)
Is able to keep a good record of what was done and why.	Notebook is not at all clear, or illegible. No clear record of anything that happened in the lab.	Activities in the lab are only presented very briefly. It would be impossible for someone else to understand what has been done.	Records have been kept but may be lacking in detail or in completeness.	Records are clear, detailed and complete.
Is able to demonstrate thinking behind their methods.	No reason provided as to why they chose the experimental method.	A description is provided which explains what they did but not why. Unclear as to why the methods might work or if they are appropriate.	An explanation is provided which explains why they chose that experimental method, but this might be lacking in detail or incomplete. Some mention of uncertainties.	A detailed explanation behind their chosen experimental method is provided and is complete. All uncertainties have been mentioned and qualitatively evaluated.
Is able to make accurate measurements.	No measurements are taken.	Only one measurement is taken, or is taken in a very poor manner.	A few measurements are taken but each measurement is vastly different to the other. Limited explanation given or there is some ambiguity.	Several measurements are taken. Clear explanation is provided that describes the differences observed in the measurements.
Is able to present results in a clear manner.	No results presented.	Results are presented with no description. Graphs might be missing axes, units might be absent.	Results are generally presented with some description but there may be some ambiguity.	Results are presented with labels, annotations and descriptions in a clear and unambiguous manner.

Is able to make sense of what has been done.	No attempt is made to place the experiment in a wider context.	Some limited sense-making is done when specifically asked to do so. No examples provided.	Context is provided. The meaning of results are discussed but perhaps only briefly or to a limited extent.	Context is provided. The meaning of results are discussed in full providing examples where necessary.
--	--	---	--	---

References

- Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K.R. (2017). Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning. *CBE-Life Sciences Education*, 16, ar56.
- Bernhard, J. (2017). Beyond active learning: Critical factors for learning in labs. In *7th Research in Engineering Education Symposium, 6-8 July 2017*. Bogotá, Colombia.
- Brookes, D. T., & Etkina, E. (2015). The importance of language in students' reasoning about heat in thermodynamic processes. *International Journal of Science Education*, 37(5-6), 759-779.
- Case, J. M. (2019). A third approach beyond the false dichotomy between teacher- and student-centred approaches in the engineering classroom. *European Journal of Engineering Education*, 44(5), 644-649.
- Eddy, S. L., & Brownell, S. E. (2016). Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. *Physical Review Physics Education Research*, 12, 020106.
- Elby, A. (2001). Helping physics students learn how to learn. *Physics Education Research, American Journal of Physics Supplement*, 69(7), S54-S64.
- Etkina, E., Murthy, S., & Zou, X. (2006). Using introductory labs to engage students in experimental design. *American Journal of Physics*, 74(11), 979-986.
- Etkina, E., & Planinšič, G. (2014, March). Thinking like a scientist, *Physics World*.
- Etkina, E., & Van Heuvelen, A. (2007). Investigative Science Learning Environment - A Science Process Approach to Learning Physics. In: E. F. Redish and P. Cooney, (Eds.), *Research Based Reform of University Physics*, (AAPT).

- Etkina, E., Van Heuvelen, A., White-Brahmia, S., Brookes, D. T., Gentile, M., Murthy, S., Rosengrant, D., & Warren, A. (2006). Scientific abilities and their assessment. *Physical Review Special Topics – Physics Education Research* 2, 020103.
- Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, 332(6064), 1213-1216
- Hake, R. R. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. (1998). *American Journal of Physics*, 66(1), 64-74.
- Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. *The Physics Teacher*, 30, 141-158.
- Lorenzo, M., Crouch, C. H., & Mazur E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics*, 74(2), 118-122.
- Mazur, E. (1997). Peer instruction: A user's manual. Upper Saddle River, NJ: Prentice Hall.
- Meyer, J. H. F., & Land, R. (2003). Threshold concepts and troublesome knowledge: Linkages to ways of thinking and practising within the disciplines. *Enhancing Teaching-Learning Environments in Undergraduate Courses Project*, (Occasional Report 4). Edinburgh, Scotland: Universities of Edinburgh.
- National Science Foundation. (2017). *Women, minorities, and persons with disabilities in science and engineering*. Arlington, VA: Author.
- Peter, M., Khoo, E., Cowie, B., Scott, J., & Round, H. (2017). *Reengineering an engineering course: How flipped classrooms afford transformative teaching, learning, and workplace competency*. Hamilton, New Zealand: University of Waikato.
- Poklinek Čančula, M., Planinšič, G., & Etkina, E. (2015). Analyzing patterns in experts' approaches to solving experimental problems. *American Journal of Physics*, 83(4), 366-374.
- Quale, A. (2011). On the role of mathematics in physics. *Science and Education* 20, 359-372.
- Redish, E. F., Saul, J. M., & Steinberg, R. N. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66(3), 212-224.

- Rylands, L. J., & Coady, C. (2009). Performance of students with weak mathematics in first-year mathematics and science. *International Journal of Mathematical Education in Science and Technology*, 40(6), 741-753.
- Salehi, S., Burkholder, E., Lepage, G. P., Pollock, S., & Wieman, C. (2019). Demographic gaps or preparation gaps?: The large impact of incoming preparation on performance of students in introductory physics. *Physical Review Physics Education Research*, 15, 020114.
- Tumarino, J. & Redish, E. F. (2007). Elements of a cognitive model of physics problem solving: Epistemic games. *Physical Review Special—Physics Education Research*, 3, 020101.
- Uttl, B., White, C. A., & Gonzalez, D. W. (2017). *Studies in Educational Evaluation* 54, 22-42.
- Wage, K. E., Buck, J.R., Wright, C. H. G & Welch, T. B. (2005). The Signals and Systems Concept Inventory. *IEEE Transactions on Education*, 48(3), 448-461.
- Wieman, C. (2015). A Better way to evaluate undergraduate teaching. *Change: The Magazine of Higher Learning*, 47(1), 6-15.
- Wilson, M. T. (2014). Student and expert perceptions of the role of mathematics within physics. *Waikato Journal of Education*, 19(2), 93–103.