NPF NATIONAL PROJECT FUND AKO AOTEAROA

Research Report

Transforming Tertiary Science Education: Improving learning during lectures

Ben Kennedy, Erik Brogt, Zoe Jordens, Alison Jolley, Rosie Bradshaw, Maggie Hartnett, Billy O'Steen, Eva Hartung, Annie Soutter, Gemma Cartwright and Natalie Burr

AKO AOTEAROA NATIONAL CENTRE FOR TERTIARY TEACHING EXCELLENCE
MASSEY UNIVERSITY TE KUNENGA KI PŪREHUROA UNIVERSITY OF NEW ZEALAND
UC UNIVERSITY OF CANTERBURY

TRANSFORMINg TERTIARy SCIENCE Edu CATION: IMPROvINg lEARNINg duRINg lECTuRES

Ben Kennedy, Erik Brogt, Zoe Jordens, Alison Jolley, Rosie Bradshaw, Maggie Hartnett, Billy O'Steen, Eva Hartung, Annie Soutter, g emma Cartwright and Natalie Burr

contents

This project was commissioned through the Ako Aotearoa National Project Fund 2010. More information is available at www.akoaotearoa.ac.nz/projects/transforming-tertiary-science-education

PuBlISHEd By Ako Aotearoa National Centre for Tertiary Teaching Excellence
PO Box 756
Wellington 6140

ISBN 978-1-927202-35-7 (print)
ISBN 978-1-927202-37-1 (online)

March 2013

Summary

Science education research shows that a traditional, stand-and-deliver lecture format is less effective than teaching strategies that are learner-centred and that promote active engagement. The Carl Wieman Science Education Initiative (CWSEI) has used this research to develop resources to improve learning in university science courses. We report on a successful adaptation and implementation of CWSEI in the New Zealand university context.

This two-year project at Massey University and the University of Canterbury began by using perception and concept surveys before and after undergraduate science courses to measure students' attitudes towards science as well as their knowledge. Using these data, and classroom observations of student engagement and corroborating focus groups, the research team worked with lecturers to create interventions to enhance student engagement and learning in those courses. Results show several positive changes related to these interventions and they suggest several recommendations for lecturers and course coordinators. The recommendations include:

1. Make learning outcomes clear, both for the lecturer and the students; this helps to cull extraneous material and scaffold student learning.
2. Use interactive activities to improve engagement, develop deeper levels of thinking, and improve learning.
3. Intentionally foster "expert-like thinking" amongst students in the first few semesters of the degree programme.
4. Be flexible because one size does not fit all and contextual events are beyond anyone's control.

In addition to these recommendations, data collected at the Canterbury site during the 2010 and 2011 earthquakes reinforced the understanding that the most carefully designed teaching innovations are subject to contextual conditions beyond the control of academics.

Introduction: Applying the CWSEI model in a New Zealand context

The Carl Wieman Science Education Initiative (CWSEI) is a multi-million dollar endeavour at the University of British Columbia (UBC) in Canada (www.cwsei.ubc.ca) and the University of Colorado in the US. Professor Wieman, a co-recipient of the 2001 Nobel Prize in physics and current science advisor to US President Barack Obama, is leading an effort to improve science learning by implementing best practices from the educational research literature.

The CWSEI methodology emphasises an evidence-based approach (Wieman, 2007, 2009, 2010) to influence changes in teaching practices through a two-stage model that involves a “baseline” and a “transformation”. In the first stage, qualitative and quantitative data to measure the existing situation are gathered to create a baseline. During the baseline phase, the learning outcomes are evaluated and corresponding changes in course content are planned. In the subsequent transformation stage, learner-centred interventions are designed (Adams and Wieman, 2011) and implemented with the aim of achieving the new learning outcomes (Beatty *et al*., 2006; Abrami *et al*., 2008; Ruiz-Primo *et al*., 2011), and data are taken to measure the effects on student learning (Hake, 1998; Deslauriers *et al*., 2011). The desired transformation occurs at a course level and may take several years of iterations until the course is fully transformed and significant improvements in learning can be measured (Wieman, 2007).

Learner-centred techniques, similar to those advocated by Weiman, have been recommended in higher education for a considerable period of time, and numerous, well-documented “off-the-shelf” techniques are available in the literature (*e.g.* Mazur, 1997; Huba & Freed, 2000; Blumberg, 2008). For example, Dr Eric Mazur (1997) has shown the effectiveness of using peer instruction in learning physics concepts. However, despite the evidence, the uptake of these approaches in university teaching has been patchy and rarely systematic (Wieman, 2009). The CWSEI model puts a strong emphasis on motivating the department and its academic staff to invest time and to drive course transformations internally (Wieman, 2007, 2009). Offering the services of an education specialist, who is also a content expert within the department, (Wieman et al., 2010) does this. Such a “knowledgeable other” (Offerdahl, in review) with credibility within the department is essential to facilitate engagement with the project, generate uptake within the academic unit, and empower staff by providing meaningful, discipline-specific teaching methods.

Despite the success of the CWSEI method in Canada and the US, its effective transference to the New Zealand educational system, which has its own academic and educational cultures and student demographics, had not been attempted prior to this project. In addition, the level of resources available to CWSEI was not available in New Zealand, all of which provided an ideal research opportunity. Thus, in 2010, a joint team from the University of Canterbury (UC) and Massey University (MU) were commissioned through Ako Aotearoa's National Project Fund to adapt and implement CWSEI in two New Zealand university settings.

The core research team consisted of academics in natural sciences, academic development, and education: Ben Kennedy (Project Leader, UC geology), Zoe Jordens (Co-Project Leader, MU biology), Rosie Bradshaw (MU biology), Maggie Hartnett (MU education), Erik Brogt (UC academic development) and Billy O'Steen (UC education). The team was augmented by several postgraduate research assistants from Massey, Canterbury and UBC to manage the day-to-day data collection and collation procedures. In addition, Kennedy, as a former postdoctoral researcher in Wieman's group, had access to people, resources and materials at UBC and, consequently, several of our colleagues at UBC provided valuable input at various stages of the project.

Following the CWSEI methodology, we aimed for a semester in which baseline data would be collected, followed by a semester of the same course with interventions implemented. Given the timetables, this meant an 18-month project from July 2010 to December 2011, or three academic semesters, as most courses run once a year. For the first semester of 2011, an accelerated programme was planned, with half of the semester providing the baseline data, and the other half providing an opportunity to apply the interventions.

In order to bring all team members up to speed with the CWSEI methodologies and resources, two three-hour virtual workshops were organised via the KAREN videolink network. The first workshop, held just prior to the baseline data semester, dealt with the reasoning behind, and the practical implementation of, examining and reconsidering course and lecture outcomes. The second workshop, run in early January 2011, focused on potential interventions and ways to implement those in the courses. Along with the core research team, several research assistants also participated in these workshops. Both workshops were based on materials from UBC and are now available on the Ako Aotearoa web space for this project.

Methodology and data sources

Target courses

In total, six courses were targeted in this study: four geology courses at UC and two biology courses at MU. The courses were chosen based on three criteria: (1) the course coordinators' willingness to participate, (2) the involvement of two different science subjects, and (3) including science courses at two different universities. Table 1 summarises the basic information for these courses. All courses were offered once per academic year. The student populations varied because first-year classes included students from other degree programmes (e.g. engineering or arts in the case of first-year geology classes) or students hoping to pursue the veterinary programme in the case of first year biology at MU. Course demographics were generally 60:40 male to female in geology classes at UC and 30:70 in biology at MU.

Of these six courses, GEOL 242 and GEOL 336 had already undergone some CWSEI modifications in previous years as Ben Kennedy was the course coordinator and a lecturer in both and similar courses exist in the geology curriculum at UBC. As a result, these courses were already partially transformed. For those and all other courses, we followed the CWSEI methodology by collecting data in both the baseline and intervention phases of the project, with the following objectives in mind:

1. Describe the local classroom situation in a qualitative manner to suggest suitable interventions.
2. Assess student engagement through observations.
3. Assess student learning in class by means of a quantitative concept test administered at the beginning and end of the course.
4. Assess student attitudes toward science by means of a quantitative perception survey at the beginning and end of the course.
5. Corroborate observation data (in particular on student engagement) with student focus group interviews (when possible).
6. Determine the efficacy of the interventions.

After the baseline semester, interventions were developed in consultation and collaboration with teaching staff of each course. Table 2 summarises the various types of data taken in each of the courses in both the baseline and intervention phases of the project.

Impact of the Canterbury earthquakes

The sequence of Canterbury earthquakes began in September 2010 and made baseline data in the geology department difficult to obtain because classes were cancelled and staff members swapped in and out of teaching responsibilities, as their expertise was required by the city and ongoing earthquake research. Then, the 22 February 2011 earthquake completely shut down the campus at Canterbury for several weeks. With the majority of the campus's infrastructure out of commission, learning environments ranged from online courses to lectures situated in makeshift tents. This meant that both students and staff had to rapidly adapt to new teaching situations. As a result, data collection for the project at UC was put on hold for the first semester of 2011. Instead, data were collected in the first semester of 2012.

This also meant that the interventions for the transformed courses of the second semester of 2011 could not be planned adequately in the first semester. Instead, greater emphasis was put on academic development, with interventions developed and implemented *in situ* in close collaboration with the lecturers during that second semester, rather than in advance.

Table 1: Basic information about the targeted courses and data gathered

Course name and number	Number of students	Number of lecturers	Semester offered
Planet Earth: An introduction to geology (GEOL 111)	~190	5	S1, 2012*
Environmental Geohazards (GEOL 113)	~90	4	S2, 2010; S2, 2011
Rocks, minerals and ores (GEOL 242)	~110	2	S2, 2012*
Magmatic systems and volcanology (GEOL 336)	~70	2	S2, 2010; S2, 2011
Biology of Cells	~500	1	S1, 2011
The Microbial World	~20	3	S2, 2010; S2, 2011

Table 2: Data taken in the courses under investigation

Course name	Teaching mode (baseline)	Baseline data collected	Interventions	Interventions data collected
Planet Earth: An introduction to geology	Traditional lecture, lab**, field trip*	Perception (pre, post), concept (pre, post), classroom observations	Data-driven academic development with teaching staff	Perception (post), concept (post)
Environmental Geohazards	Traditional lecture, workshop, field trip	Perception (pre, post), concept (pre, post), classroom observations, focus group interviews	Data-driven academic development with teaching staff and reformulation of course outcomes	Perception (pre, post), concept (pre, post), classroom observations, focus group interview, online survey
Rocks, minerals and ores	Interactive lecture, lab**	Perception (pre, post), concept (pre, post), classroom observations	Increased use of clickers and in-class exercises	Perception (pre, post), concept (pre, post), classroom observations
Magmatic systems and volcanology	Interactive lecture, lab**	Perception (pre, post), concept (pre, post), classroom observations, focus group interviews	Increased use of clickers and in-class exercises	Perception (pre, post), concept (pre, post), classroom observations, focus group interviews, online survey
Biology of Cells	Traditional lecture, some interactive exercises, lab**	Perception (pre, post), concept (pre, post), classroom observations	Reformulation of learning outcomes, reduction of content, use of clickers	Perception (pre, post), concept (pre, post), classroom observations
The Microbial World	Traditional lecture, lab**	Perception (pre, post), concept (post), classroom observations, student focus group interviews	Reformulation of learning outcomes, reduction of content, use of clickers and in-class exercises	Perception (pre, post), concept (pre, post), classroom observations, student focus group interviews, lecturer interviews

*Generally, course transformation was focused on lectures. * Denotes a course originally planned to be offered in 2011, but postponed as a result of the 22 February 2011 Christchurch earthquake. ** indicates parts of courses that were not changed from previous iterations of the course.*

Instruments

To collect robust and triangulated baseline and intervention data, we used a mixed-methods research design, utilising qualitative and quantitative instruments and approaches including pre- and post-perception surveys, pre- and post-concept tests, classroom observations, and focus group interviews. These instruments and approaches are carefully designed (Adams and Wieman, 2011) to measure and corroborate (amongst themselves) students' changes in expert-like thinking (Wieman, 2009), concept knowledge (Hake, 1998), classroom engagement (Lane *et al*., in prep) and opinion (Merton *et al*., 1990; Cooper & Schindler, 2011) during a defined period of time (in this case, an academic semester).

Perception surveys

A validated survey to measure student perceptions of science was developed by UBC (Jolley *et al*., 2012) containing 30 five-point Likert scale statements. UC adapted this instrument slightly to reflect our local situations (*e.g.* by changing "earth and ocean science" to "geological science") and administered it at the start and finish of the baseline and transformation courses. Similarly, for the biology papers at MU, a survey developed by UBC containing 40 five-point Likert scale statements was modified slightly for the local situation. The survey measures student responses on a novice–expert continuum and illustrates how their perceptions change after taking the course. Student responses were then grouped into categories and analysed to characterise more general aspects of student perceptions (Jolley *et al*., 2012).

Concept tests

No validated instruments exist to accurately measure students' content knowledge in geology and biology for the courses under investigation. Either the instrument does not exist at all or those that do exist, such as the Geoscience Concept Inventory (GCI) by Libarkin and Anderson (2005), only partially apply to the curricula. Developing fully validated concept tests was beyond the scope of this project, as that is typically a multi-year, iterative development and validation process. Instead, we used questions from UBC where possible, especially in geology because similar courses exist at UBC. In addition, we used the baseline classroom observations in GEOL113 to infer learning outcomes and develop assessment questions to measure those outcomes. These assessment questions then became draft concept test questions and, where possible, were integrated with existing questions from the GCI (Libarkin & Anderson, 2005) or UBC. All questions for all of the courses were informally validated by multiple lecturers and postgraduate students. However, we acknowledge that these tests are not fully validated yet.

Classroom observations

Lincoln and Guba's (1985) Naturalistic Inquiry framed our approach to observations and desire to create a "thick description" case study of what is happening inside the classrooms at MU and UC. Because teaching and learning are very context-dependent, it is important to get an understanding of the socio-dynamics of the classroom, classroom climate, teaching style, interactions between teacher and students, and student and teacher (non) verbal behaviours. This understanding of the dynamics and personalities involved is vital to create interventions that are well aligned with the teaching context so that they are authentic for the teacher, suit her or his teaching style, are pedagogically meaningful within the particular teaching situation, and are engaging for the students.

Covert classroom observations provided an indication of student engagement. We followed an observation protocol currently in development at UBC (Lane & Harris, in prep) and that of Hora and Ferrare (2009) as the basis for our own observation protocols. Observations were generally done by research assistants with the following procedure:

1. Before each lecture began, 10 students were selected at random. Care was taken not to select students sitting directly beside each other in order to minimise the effect of students distracting the person immediately beside them.
2. The observer almost always sat in the same area of the classroom, in order to minimise variation throughout the semester (for a discussion of this effect see Lane and Harris, in prep).

In several instances, multiple observers were used to both observe different students in a single class and to observe the same 10 students. The results from the observations showed "inter-rater reliability" with engagement data between observers and with different student groups, giving us confidence that observing a small sample of the class could be used reliably as a proxy for the class as a whole. These observations could also be linked to teaching behaviours that appeared to engage, or disengage, students and could be targeted for retention or elimination, as appropriate.

One special case was GEOL113: Environmental Geohazards at UC. During the baseline semester, this course was observed by two assistants, one with a background in geology and one with a background in education. The purpose was not only to monitor student engagement and note opportunities for suitable interventions, but also to infer learning outcomes and develop questions that could be used for the first iteration of a concept test.

Focus g roups

At UC and MU, focus groups (Merton *et al*., 1990; Cooper & Schindler, 2011) consisting of a "purposive sample" (Lincoln and Guba, 1985) of students who responded to invitations to attend in each of the geology courses were conducted with two intentions. Firstly, the focus groups were used toward the end of the data collection from the observations with the aim of confirming, refuting, and adapting the emerging grounded theories about what engaged students during lectures. We wanted to validate our interpretation of the observational data with students who had witnessed the same phenomena. Secondly, the focus groups were used to collect additional data beyond our observations and the surveys that would enable us to produce "thicker descriptions" (Lincoln and Guba, 1985) of our findings. This became especially important when we reported back to the lecturers of the courses, in that we could provide actual quotations from students regarding specific aspects of the teaching and learning that they found engaging or disengaging. It is important to note that the findings from these focus groups, because of their invitational and self-selected constitution, present important limitations to be considered.

Interventions

Clari cation and communication of learning outcomes

A central component of the CWSEI model is to examine, and re-examine, the learning outcomes, both for the course in its totality and for each individual lecture/ lab/tutorial. To do so, lecturers were invited to consider using the revised Bloom's taxonomy (*e.g.* Anderson & Krathwohl, 2001) as a framework for evaluating their learning outcomes. This is often helpful in determining how course materials align, or not, with the intended outcomes, and to determine if there are gaps in the curriculum that would make certain learning outcomes difficult to achieve.

Another objective of examining the learning outcomes was to increase the explicit congruence and articulation of the course outcomes and the lecture outcomes during lectures. For example, in GEOL242, the following congruent course and lecture outcomes were developed and clearly articulated during a lecture:

s ! FERTHIS **course**, students should be able to:
  ° Apply the relevant concepts of chemistry and physics to explain mineralogic, igneous, metamorphic and ore-forming processes using examples from New Zealand and the rest of the world.

s ! FERTHIS **lecture**, students should be able to:
  ° Use the chemical properties shown in the periodic table (*i.e.* atomic radii, charges) to:
    A] identify different parts of a mineral formula (*i.e.* cation)
    B] represent the size of the ions
    C] demonstrate balance between anions and cations within minerals
    D] label parts of the chemical formula that allow minerals to be classified.

Several techniques were used to accomplish this. Learning outcomes were more explicitly communicated to students during lectures (typically by means of a slide at the start of the presentation explaining lecture learning outcomes) and they were more frequently and explicitly referred to (*e.g.* Simon and Taylor, 2009). In addition, transitions between individual lectures were more explicitly linked in order to scaffold the learning and to increase students' clarity on how different aspects of the lecture course and assessment are tied together (Biggs & Tang, 2007). The aim was to ensure that students know what to expect and why certain material is covered (Marsh, 2007). The use of learning outcomes assisted staff with culling extraneous material and reducing the students' cognitive load by allowing them to focus more on the tasks themselves.

In-class exercises

In-class exercises were used in a variety of ways and forms: as feedback tools, formative assessment, and summative assessment.

As a feedback tool, the in-class exercises were often quite short (*e.g.* "one-minute papers", Angelo & Cross, 1993) such as asking students to write what they had understood and what they were unclear about at the end of the lecture. This could then be used for a "Just In

Time" teaching scenario (*e.g.* Novak *et al*., 1999), where the lecturer could adapt the next lecture based on the "muddiest point" students were still struggling with.

The formative assessment form of in-class exercises fulfilled a similar feedback role, but these exercises were typically more elaborate, requiring more writing and drawing. They were primarily designed to verify if students had understood concepts or were able to integrate concepts and transfer them to a new domain. In GEOL242 and GEOL336 at UC, these exercises were discussed extensively in the subsequent lectures. Additionally, in some cases the in-class exercises took the whole lecture and then became part of the marked assessment for the course.

Clickers and peer discussion

Electronic responder devices (commonly referred to as "clickers") have been used for a number of years in a variety of pedagogical applications. For this project, the clickers were used as a formative assessment tool (*e.g.* Mazur, 1997) where students answered a multiple-choice question usually with prior peer discussion (*e.g.* Deslaurier *et al*., 2011) and then the lecturer could choose to either discuss the topic further or guide students to discuss it within small groups.

Academic development

At UC, the Academic Development Group (ADG) offers support for staff on matters related to teaching techniques, pedagogy, curriculum, and assessment design. Erik Brogt from the ADG observed many lectures in his capacity as an academic developer and provided individual feedback on teaching to lecturers. Much of the feedback focused around being attuned to the classroom environment and being responsive to learner-centred behaviour, in particular around questioning techniques to engage students better. In addition, he collaborated with individual lecturers to discuss content representation and the structure or flow of lectures in relation to the defined learning outcomes of the lecture.

At MU, the education advisor, Maggie Hartnett, an experienced classroom observer and lecturer in education, also provided similar feedback to lecturers and helped them in aligning their teaching activities with the desired learning outcomes.

Results

Results at Massey u niversity

At Massey University, an analysis of the baseline classroom observations of the second-year microbiology paper led to a categorisation of teaching practices into a number of themes (Figure 1) and identified *skimming or skipping material* as the most significant factor that disengaged students. In response to this, the amount of content was reduced by approximately 10 *per cent* in the intervention period to avoid the need to skim over material due to time constraints. Conversely, mean student engagement was highest for *open questions, multimedia*, and *clarifications and expectations*, and these were therefore increased in the intervention period. Other interventions included enhanced (and increased use of) learning outcomes, the introduction of clicker questions designed to promote peer-to-peer discussion, and in-class activities. Classroom observations, matched for each lecture, showed increased student engagement after the introduction of interventions for all lecturers (Figure 2) in the transformed year when compared to the baseline.

Data showing the relationship between student engagement and teaching practice categories (Figure 1) and the relative number of instances spent on each category (Figure 3) were used for individual staff academic development at the end of the baseline period and again at the end of the study. These highlighted changes in the structure of lecture sessions before and after the introduction of interventions enabling lecturers to see how their teaching had changed and student engagement had increased.

At Massey University, an analysis of the baseline classroom observations of the second-year microbiology paper led to a categorisation of teaching practices into a number of themes (Figure 1) and identified skimming or skipping material as the most significant factor that disengaged students. In response to this, the amount of content was reduced by approximately 10 per cent in the intervention period.

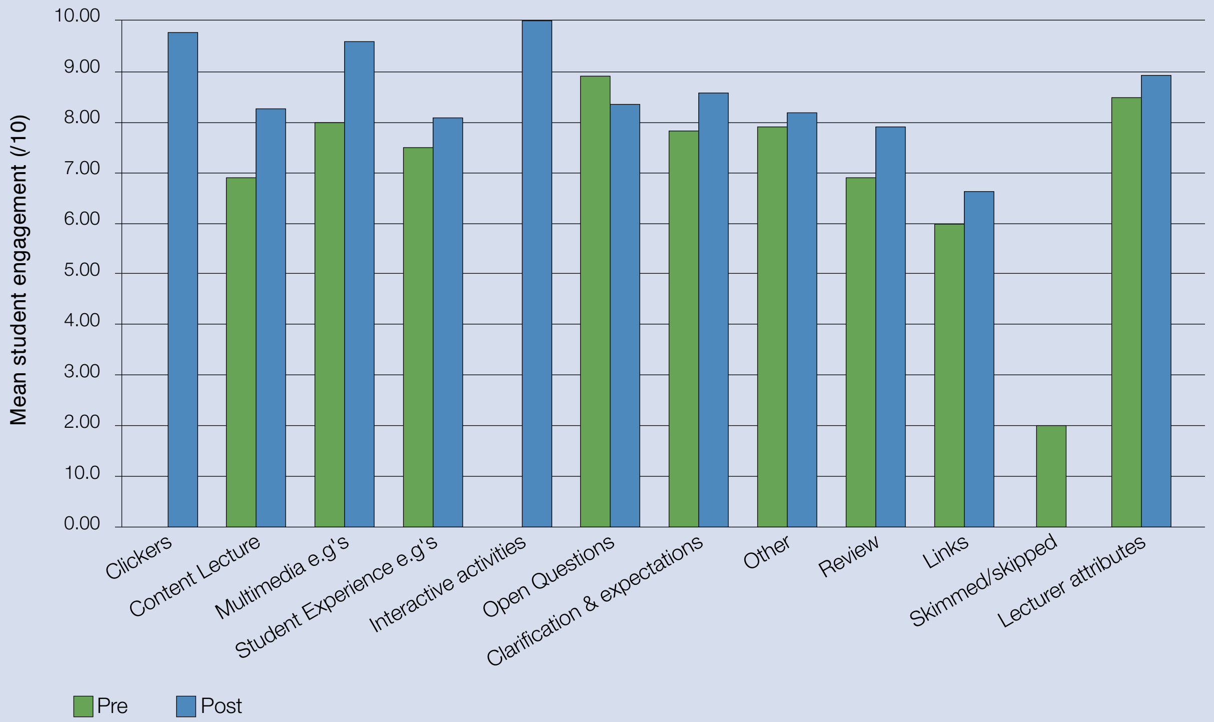


Figure 1: Mean student engagement in microbiology classes with ZJ showing engagement varied with category of teaching activity (pre, baseline period; post, intervention period)

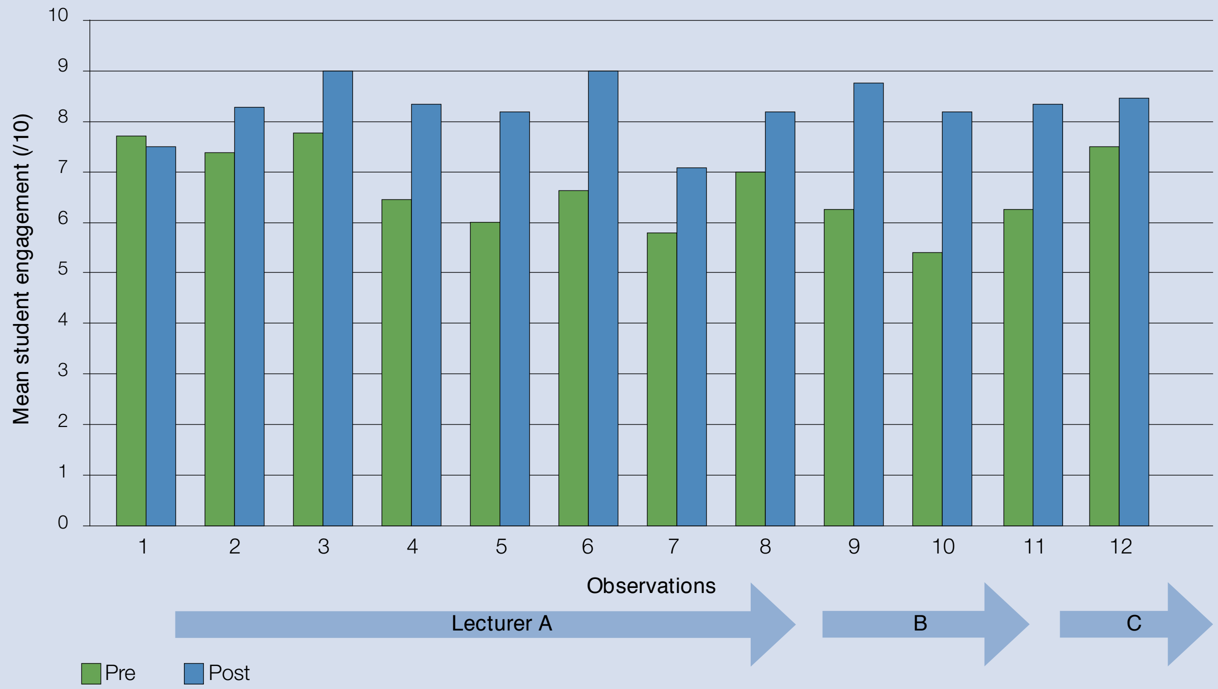
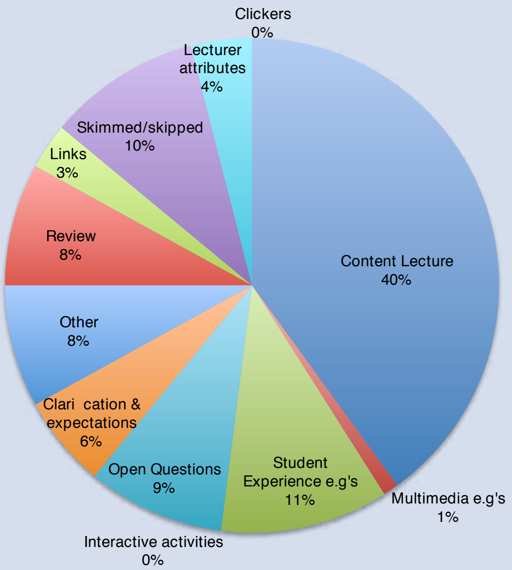


Figure 2: Mean student engagement per lecture in the baseline microbiology course in 2010, compared to the same lectures in the transformed microbiology course in 2011 (pre= baseline period; post = intervention period)

Figure 3 compares the activities of one of the lecturers in the baseline and the transformed courses. The amount of content given in traditional lecture format in lectures was substantially reduced in the intervention stage (from 40 *per cent* to 15 *per cent* of teaching instances) in favour of learner-centred activities like clicker questions or other activities. In addition, lectures were reconfigured such that no material needed to be skipped or skimmed, which was one of the activities seen to disengage students (Figure 1).

Although the microbiology class was small (and hence the very small numbers of matched students in this study), the analysis of students' responses on the perception and concept surveys showed some positive trends that were similar to the data collected in larger courses at Massey and Canterbury. Overall, the perceptions during the transformed course shifted towards those associated with biology experts (12.3% shift) in contrast to the baseline course where attitudes did not shift towards experts (overall shift away -2.9%), with the most positive shift in problem solving (Table 3).



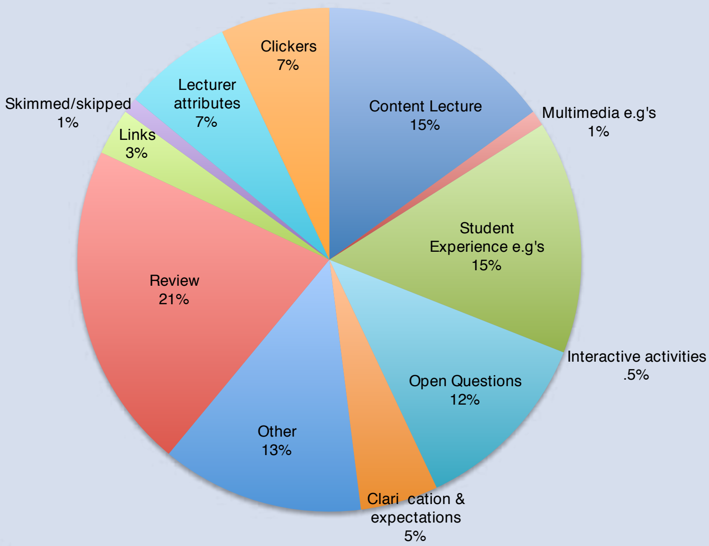


Figure 3: to left: Relative number of instances spent on each teaching category in the microbiology course by lecturer ZJ (baseline); to right: Relative number of instances spent on each teaching category in the microbiology course by lecturer ZJ (transformed course)

Table 3: Perception survey changes in the baseline and transformed microbiology course at Massey University

	Baseline 2010 % shift during course	Transformed 2011 % shift during course
Overall (7 categories)	-2.9*	12.3
'Real world connection'	1	3.3
Problem solving (difficulty)	-14.8*	23.6

** Negative result indicates a shift away from "experts". Full list of questions can be found at http://akoaotearoa.ac.nz/projects/transforming-tertiary-science-education*

Table 4: Concept test and semester test results for the microbiology course

	Pre-course Concept Test mean (sd) %	End of Course Concept Test mean (sd) %	Semester Test mean (sd) %	Final Exam mean (sd) %
2010 *(n=12)*	Not determined*	48 (15)	54 (21)	57 (14)
2011 *(n=8)*	40 (10)	64 (15)	63 (20)	64 (16)

** The original concept test proved too easy and was re-written for subsequent use.*
The small n *reflects the matched surveys pre-post.*

Importantly, the transformed course showed improved learning as measured by the concept test, semester test, and final exam marks, with mean increases of 16 *per cent*, 9 *per cent* and 7 *per cent* respectively (Table 4). This supports the notion that increased engagement is associated with increased learning.

Transformation of the large, first-year Cell Biology course occurred within a single semester. The concept tests indicated slight gains in the transformed part of the semester (statistically significant in top-performing students, and also middle-performing students when the two most difficult concept survey questions were removed). This indicated that the concept test was too difficult, and required further validation. Clicker questions resulted in a high level of student engagement (mean 9.6/10) compared to the mean of other activities. Clickers also were very popular with students: an end-of-course vote showed 95 *per cent* of students wanted to continue using them. Full details of cell biology course data can be retrieved from http://akoaotearoa.ac.nz/projects/transforming-tertiary-science-education

Transformed courses showed improved learning as measured by the concept test, semester test, and final exam marks, with mean increases of up to 16 per cent. This supports the notion that increased engagement is associated with increased learning.

For the first-year biology course, overall student engagement dropped in the transformed course during the second half of the semester. The perception-survey measures of student learning showed negative shifts in student perceptions relative to 'biology experts' during the semester. We attribute the lack of improvement in student perceptions and engagement in the transformed part of the semester to: (1) a general negative shift in student perceptions over the course of their first semester at university (Adams *et al*., 2006); (2) different material was taught in the second half of the semester compared to the first; and (3) the interactive nature of this course prior to interventions, with only 16 *per cent* of time spent in traditional lecture format (reducing to 12 *per cent* with interventions). However, encouragingly, the same transformed course in the following year (2012) continued the positive trend of improved exam results related to the transformed course.

Final interviews with each of the four lecturers involved at Massey University explored their responses to the study's approach to enhancing teaching and learning. The main benefits of taking part in the study were perceived to be: (1) the opportunity for improvement (in teaching) and enhanced student learning; (2) authentic and meaningful professional development; (3) credibility and influence (being involved in a study and collecting data); and (4) more interactive teaching. Lecturers highlighted the need for two key resources: time and access to expertise. The challenges noted included content constraints, ensuring conceptual understanding, and the reliability of the data. Lecturers perceived the strengths of the approach to be enhancing student interaction and interest, and changing their teaching. A selection of quotes from lecturers that illustrate the benefits of taking part in the study is given below:

- Enhanced student learning:
  Obviously [to] enhance the student learning is the absolute biggest thing, the most important thing and take it to as wide a staff audience as possible. (Lecturer)

- Authentic and meaningful professional development:
  Much better, much, much better, because you've actually been taking it into the classroom you're teaching in, it's actually for your context. You're doing it and using it, so it's meaningful, it's not just some theoretical situation. Everything has been applied, everything we've thought about, discussed has been for a particular situation, lecture that I'm doing and so it's all relevant, so all of the time that I've put in has been worth it. I suppose the point I'm trying to make is it's all been [an] efficient use of my time because it's all been put straight into practice. (Lecturer)

- More interactive teaching:
  On the clickers: I don't think I really understood at the start that they were merely a tool to start peer to peer discussion. I thought they were just a multi-choice test. Once I realised that they started the interaction and they are effectively a tool for starting small group discussion and it's that, that enhances their learning. I will now think much more carefully how I address; in fact towards the end

this became obvious in that I wrote the questions accordingly. It was a learning process for me obviously, so now I will be far more aware of the fact that I am using them to promote discussion. (Lecturer)

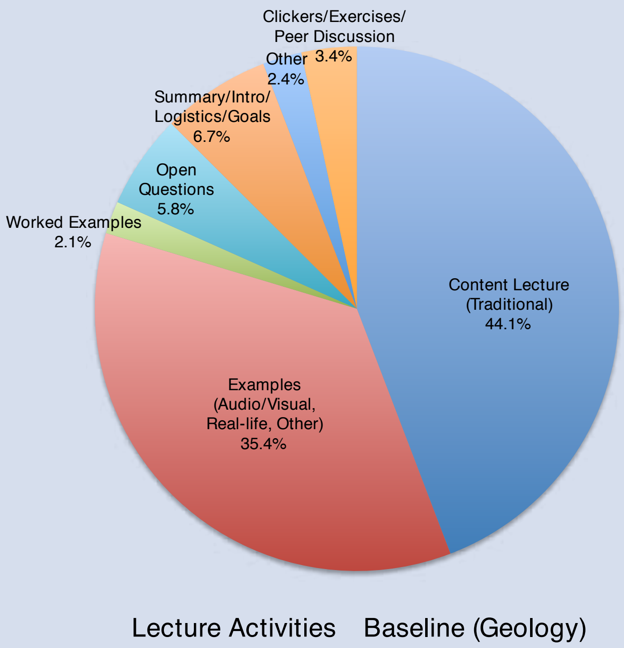
I think it encouraged me to actually put more emphasis on interactive approaches. (Lecturer)

Summary of key ndings at Massey University

1. Classroom observations of teaching and student engagement were reported to be very useful for lecturers' academic development.
2. Improving learning goals, reducing content by 10 per cent and increasing student interaction (clickers) increased engagement and enhanced learning (increased exam marks).
3. Clickers were popular with students and increased student engagement and learning.
4. Increased lecturer con dence with clickers led to increased use, with more peer-to-peer discussion, further enhancing engagement.
5. Word of mouth spread from lecturers and students is encouraging other lecturers to try clickers; software has been requested in Massey lecture theatres for 2012, which will make them more accessible.

Results at the u niversity of Canterbury

The September 2010 and February 2011 earthquakes hampered data collection at UC. As a result, the CWSEI model of comparing a course's baseline data to the transformed version of the same course had to be abandoned. However, two courses – GEOL242 and GEOL336, which had already undergone partial transformation – continued to be transformed in the moment and were thus classified as transformed. These could be compared to GEOL111 and GEOL113, which underwent data-driven academic development while still trying to respond to the baseline data, with substantial transformation planned for the following semester. The differences in teaching activities between the baseline and transformed courses are shown in Figure 4. The main differences between the two types of courses were the increased use of peer discussion, in-class exercises and clicker questions in the transformed course.



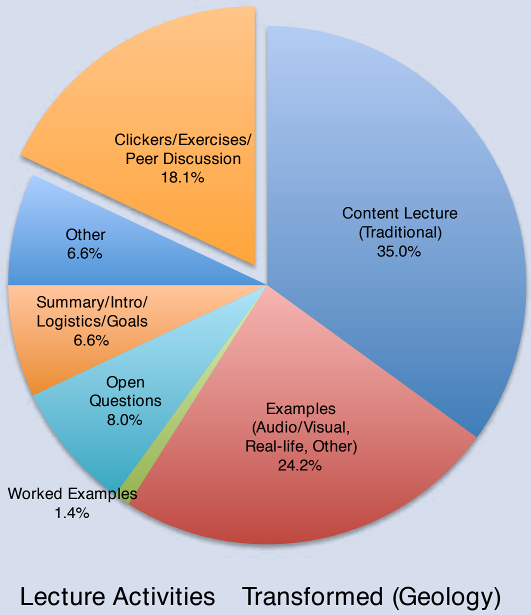


Figure 4: to left: Average percentage distribution of types of activities in the GEOL111 and GEOL113 baseline courses; and to right: Average percentage distribution of types of activities in the GEOL336 and GEOL242 transformed courses. Note: data shows instances, not total time spent.

Figure 5: Average engagement as a function of time, averaged over the transformed and baseline geology courses at University of Canterbury. Note the steady level of engagement in transformed courses.

As can be seen in Figure 5, the geology courses at UC had a high level of engagement. This is perhaps not too surprising given the increased interest in geology as a result of the earthquakes. Interestingly, the baseline courses, GEOL113 and GEOL111, both covered earthquakes as part of the curricula whereas that topic was absent in the (more advanced) transformed courses. Even so, the transformed courses showed higher engagement and particularly showed a decreased drop in engagement as a function of time. However, the effect of the difference in student population (first-year versus higher-year students) cannot be ruled out.

Table 5 shows how teacher activity correlates with student engagement. For example, when teaching activities, such as in-class exercise and peer discussion, were used in the baseline courses, high student engagement was comparable to that in the transformed courses.

However, in the baseline courses comparatively few lecture hours were dedicated to these activities and engagement dropped when these activities were not used, especially towards the end of the lecture (Figure 5).

Table 5: Average engagement per activity for all geology courses at University of Canterbury

Activity	Baseline		Transformed	
Lecture activity	*Average engagement % (400 students)*	*Sum of observed individual student hours*	*Average engagement % (240 students)*	*Sum of observed individual students hours*
Clickers/ Exercises/ Peer discussion	92	63	93	205
Content lecture (traditional)	86	809	92	396
Examples (audio/visual, real-life, other)	86	649	93	274
Worked examples	85	39	97	16
Open questions	85	106	95	91
Summary/ Intro/ Logistics/ Outcomes	77	124	84	75
Other	82	44	93	75

The sum of observed individual student hours is the total time that observed students spent doing particular activities.

Table 6: Pre- and post-course concept test (CT) scores for courses in geology at University of Canterbury

	GEOL111	2012		GEOL113	2011	
	Pre CT	*Post CT*	*Gain*	*Pre CT*	*Post CT*	*Gain*
Count (n=)	183	187	144	86	77	45
Average (%)	**39.69**	**63.21**	**0.39**	**46.66**	**57.58**	**0.20**
Stdev (%)	15.56	14.33		12.55	13.21	
	GEOL242	2012		GEOL336	2011	
	Pre CT	*Post CT*	*Gain*	*Pre CT*	*Post CT*	*Gain*
Count (n=)	107	72	58	49	50	43
Average (%)	**23.51**	**62.17**	**0.51**	**47.91**	**58.95**	**0.21**
Stdev (%)	12.54	14.32		13.24	12.65	

Pre- and post-concept tests were run in all of the courses. All classes showed that students performed better on the test at the end of the semester than when they took it at the beginning of the semester. The concepts tests run in 2011 experienced some developmental problems. For example, insufficient time was allowed for the post-test in 2011 for GEOL336, and several questions in the GEOL113 concept test were not considered to be well validated upon reflection. However, in 2012, more care was taken with the validation and timing of the tests. Learning gains were calculated according to Hake (1998). We did not compare the relative learning gains between classes as the difficulty and situations of the tests vary.

As an example, we present the results from the concept test in the transformed GEOL242 class (Figure 6). These results show the distribution of pre-concept test scores and post-concept test scores for students who took both tests. All individual students shifted to higher scores at the end of the semester compared to the beginning. The average mark improved from 23.5 *per* cent to 62.2 *per cent*, suggesting that the level of the concept test was appropriate and that students learned concepts successfully in the transformed course.

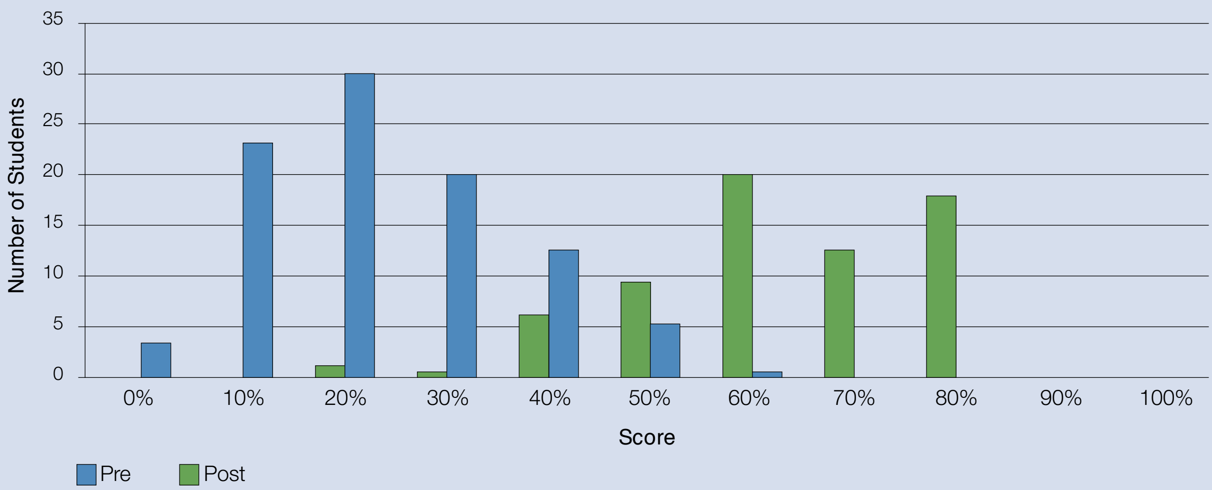


Figure 6: Pre- versus post-concept test scores in GEOL242

None of the overall shifts in student perceptions were statistically significant, as transformed courses did not affect students' perceptions more than baseline courses (Appendix 1). However, the lack of negative shifts is worth noting in comparison to physics courses (Adams *et al*., 2006). Some significant shifts could be seen in certain categories within a semester. Most significant is the shift towards expert perceptions in "the conceptual problem solving".

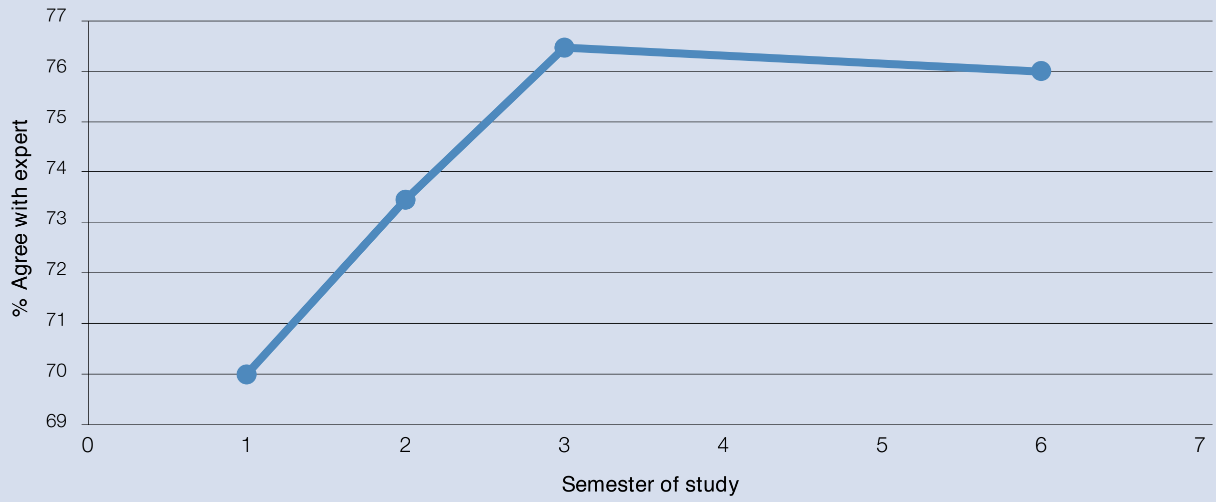


Figure 7: Mean perception score as a function of semester in study in geology at University of Canterbury

Geology is a three-year programme with two semesters per year and courses can be arranged by semester of study in order to look for changes as a student progresses through their degree. To do this, multiple years and pre- and post-course data were all averaged to obtain one overall number that represented an average student's perception within that semester. The results show that geology majors do shift towards a more expert-like perception as they progress through their degree. This shift occurs during semesters one to three with little apparent change occurring between semester three and six (Figure 7).

Student interviews and questionnaires also revealed support for the use of clickers, peer discussion, in-class exercises and formative feedback. In one course students were asked what important concepts they had learned. Of 135 responses, 105 students chose concepts that were reinforced by clicker questions, feedback from in-class exercises and/or discussion with peers. In addition, when asked at what point the concepts became clear ("clicked") for them, 70 *per cent* of students listed activities that were related to the transformed nature of the course (Figure 8).

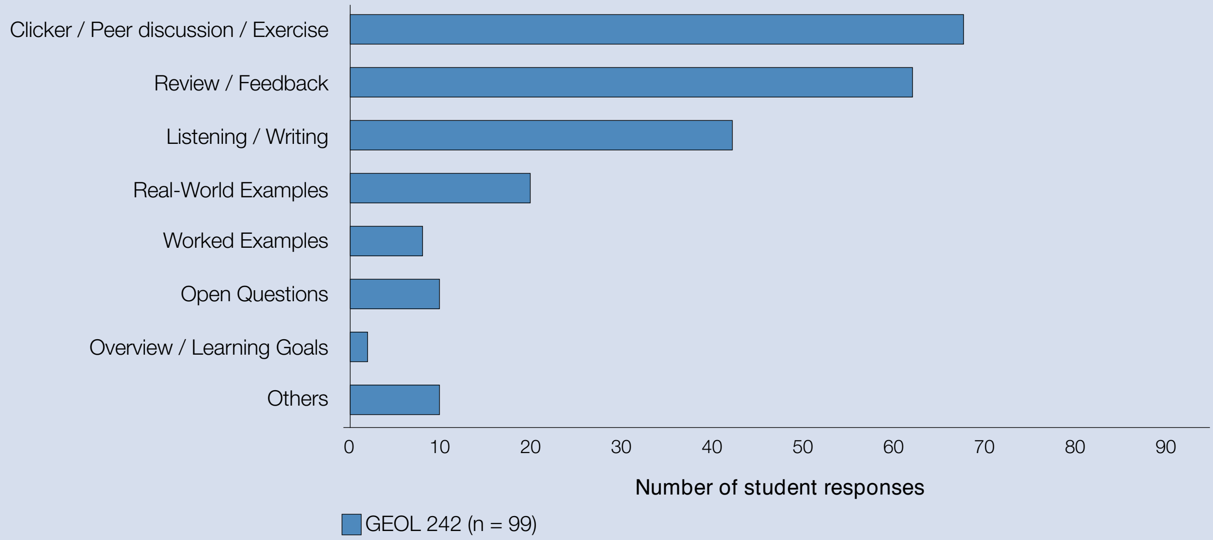


Figure 8: Written student responses to the question "At what point in the lecture did the concept 'click'?", collated as a function of classroom activity

Student interview data corroborated these findings, as exemplified by the following representative quotations:

> *I mean, I think it's more challenging in the respect that you have to engage during the class, so you can't just, like, sit there just looking cool. You actually have to think and process things.* (Student, GEOL336)

> *And you're actually, you know, processing that information as you learn it rather than just putting it all in the notes and forgetting…* (Student, GEOL336)

> *These 'alternative' teaching methods would be helpful in many other classes such as biology and other geology classes. It makes it much more enjoyable going to lectures.* (Student, GEOL336)

> *Asking us questions forces us to think even when we don't answer.* (Student, GEOL336)

The earthquake effect

The Canterbury earthquakes gave us a unique opportunity to examine student perceptions pre- and post-earthquake. In the GEOL113 course, the first perception pre-test was taken in July 2010, prior to any earthquakes. The perception pre-test for the next iteration of the course was in July 2011, after numerous earthquakes. An example of the potential earthquake effect is the response to the statement: "Geological Science has little relation to what I experience in the real world". Before the earthquakes, 72 *per cent* of students disagreed with this statement and after the earthquakes, 97 *per cent* of students disagreed with this statement.

Summary of key findings at University of Canterbury

1. Overall student engagement was very high in Canterbury geology classes.
2. Baseline courses with less interactive Wieman-approved activities had lower levels of student engagement than transformed courses.
3. Concept tests showed that student learning within a semester was comparable to transformed physics courses in the US (Hake, 1998).
4. Students within geology already had highly expert-like perceptions that showed most significant improvements in the first half of their geology degrees.
5. Student interviews and questionnaires showed unanimous support for the use of interactive activities and peer discussion.

Discussion

We successfully used and adapted the CWSEI model from UBC to initiate and sustain changes in teaching practices in the Geology department at UC and the Institute of Molecular Biosciences at MU. Specifically, we developed and implemented tools to measure student engagement and learning in our classrooms and to use this data to inform further changes. Our results show that the CWSEI model improved student engagement and learning in the New Zealand context. The improved student engagement and learning was supported by follow-up student focus groups and questionnaires.

Implications for the classroom

At both UC and MU, learning outcomes were successfully used to streamline teaching and assessment and to inform students of course expectations. Learning outcome-aligned interactive techniques in the transformed courses at both MU and UC (such as in-class exercises and clicker questions with peer discussion) correlated with apparently high student engagement (Figure 1 and Table 5). This was exemplified by The Microbial World at Massey, which was the only course where a true baseline could be compared to a fully transformed course. The transformed course showed increases in student engagement in almost every class of the semester and even with different lecturers (Figure 2). These changes in student engagement correlate with the changes in lecturer activity illustrated in Figure 3. At UC the transformed courses with more interactive learning strategies (Figure 5) show higher engagement when compared to the more traditional baseline courses. Student interviews and questionnaires validated that student engagement in transformed classes, when compared to the baseline courses, involved “having to think”, “processing information”, “applying more”, “practising more”, and “not just forgetting”. Where concept tests were sufficiently validated and suitably measured, our transformed courses compare well to transformed courses in the US (Hake, 1998).

Implications for the programme

Our perception survey results are consistent with UBC (Jolley, pers. comm), and rarely show significant shifts within a semester, even in a fully transformed course. However, student perceptions do show an increase in the direction of expert thinking over three years of university study. This particularly highlights the finding that these changes seem to occur between semester one and three. This emphasises the importance of first-year courses in developing expert-like thinking in students at an early stage in their degree.

At UC significant improvements could be seen in the category of Conceptual Problem Solving; however, very little improvement was seen in categories such as Science and Society or Human Science Interaction. One interesting feature is the difference in perception data before and after the earthquake, which will be the subject of future study.

Implications for department, college, and university

This project was possible because team members consisted of academics and research assistants from within the science departments, who worked with education specialists and academic developers. The major hurdles for projects like this are the large commitment of time and complications in organising data collection and faculty availability. However, time was saved by using the existing resources from the Wieman group and spreading the time load between the group members. The use of numerical observational data, measures of learning gains, student interviews, traditional academic development and college-level support provided sufficient momentum with which to engage academics.

Recommendations, lessons learned and future work

Based on the findings presented here, the following broad recommendations can be made:

1. *Make learning outcomes clear, both for the lecturer and the students; this helps to cull extraneous material and scaffold student learning.*
2. *Use interactive activities to improve engagement, develop deeper levels of thinking, and improve learning.*
3. *Intentionally foster "expert-like thinking" amongst students in the first few semesters of the degree programme.*
4. *Be flexible because one size does not fit all and contextual events, such as earthquakes, are beyond anyone's control.*

The earthquakes in Christchurch resulted in the UC project not having the clear baseline and transformed design as originally desired. However, this situation of not being able to execute a clean "true experiment" design in these sorts of projects is the norm rather than the exception. More often than not, reform projects are not constrained by forces of nature, or unwillingness on part of the teaching staff, but simply by logistics, in particular a lack of time. While UC was an extreme example, our experiences reinforce the notion that sustained reform requires a long-term process with a series of small incremental changes, each of which are readily achievable within the resource-limited teaching environment.

Sustained development

In any teaching development project, the largest risk is always that the reforms cannot be sustained after grant funding has ceased. While CWSEI is a multi-million dollar endeavour, this project had a much smaller budget from Ako Aotearoa, of which much had to be allocated for staff buy-out. In any realistic setting in which structural reform in education is addressed, resources will be limited with the most important resource not necessarily being money, but time. We addressed this issue in several ways (though admittedly partly by trial and error) and advise the following approach:

1. Invest in a (postgraduate) research assistant who can spend a significant amount of time on the project. In the latter part of the project, our research assistants (all with interest and skills in disciplinary education research) worked at least half-time, and were responsible for observations, data management and analysis. Multiple assistants work as well, but having a single lead or person responsible for all data reduces the overhead (and thus staff time) considerably.
2. Support from the university central administration or college may be available in the form of seed or teaching grants.

3. Set a limited number of goals to avoid spreading resources too thin to do anything meaningful with any one course.
4. Involve the teaching support group(s) on campus, if any. While these people generally are not disciplinary experts, they bring a wealth of experience about (measuring) tertiary teaching and learning and may have valuable connections on and off campus to draw on. We found these interactions to be very fruitful in helping us to see the project from different angles and from different academic cultural perspectives.
5. Involve departmental and college leadership.

The second risk, related to the first, is that the momentum for reform ceases with the grant. In this project, we have endeavoured to mitigate this risk through a variety of means. The most important strategy was the empowering of academic staff through academic development support on matters related to teaching and learning, pedagogy, assessment and course design to provide them with the tools necessary to sustain the reforms and teaching techniques on their own. In addition, the emphasis was on an evolutionary, rather than a revolutionary, reform for several of the courses at Canterbury. This means that the reforms become part of the standard teaching practice, instead of reforms just for the duration of the project. The relationships developed between teaching staff and academic developers as a result of this project are still utilised and will allow for further collaborations between lecturers and academic developers and bring the results of the project to the wider university community.

A significant spin-off of this research project is the formal establishment of a Geoscience Education Research Group at the University of Canterbury, housed within the Department of Geological Sciences, and staffed by both research team members and interested teaching staff from the department. In addition, the Department formally identified Geoscience Education as one of its strategic research foci for the future. Lastly, the university provided a Teaching Development Grant to develop several in-class exercises around several topics in Geohazards. These exercises are currently being field tested both at UBC and at Canterbury.

References

Abrami, P.C., Bernard, R.M., Borokhovski, E., Wade, A. Surkes, M.A., Tamim, R. & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 Meta-analysis. *Review of Educational Research*, *78*(4), 1102-1134.

Adams, W.K., Perkins, K.K., Podolefsky, N.S., Dubson, M., Finkelstein, N.D. & Weiman C.E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Phys. Rev. ST Physics Ed. Research*, *2*, 1-14.

Adams, W.K., & Wieman, C.E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education*, *33*(9), 1289-1312.

Anderson, L.W., & Krathwohl, D.R. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives*. Complete edition. New York: Longman.

Angelo, T.A., & Cross, K.P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.

Beatty, I.D., Gerace, W.J., Leonard, W.J., & Dufresne, R.J. (2006). Designing effective questions for classroom response system teaching. *Am. J. Phys*., *74*(1), 31-39.

Biggs, J., & Tang, C. (2007). *Teaching for Quality Learning at University*. Maidenhead: Open University Press/McGraw Hill.

Blumberg, P. (2008). Developing learner-centered teaching: A practical guide for faculty. San Francisco: Jossey-Bass.

Cooper, D.R., & Schindler, P.S. (2011). *Business Research Methods*. New York: McGraw-Hill.

Deslaurier, L., Wieman C., & Schelew, E. (2011). Improved Learning in a Large-Enrollment Physics Class. *Science*, *332*(6031), 862-864.

Hake, R.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.

Hora, M., & Ferrare, J. (2009). *Structured observation protocol for instruction in Institutions of Higher Education (IHEs)*. Madison, WI: University of Wisconsin-Madison, Wisconsin Center for Education Research.

Huba, M.E., & Freed, J.E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Needham Heights, MA: Allyn & Bacon.

Jolley, A., Lane, E., Kennedy, B., & Frappé-Sénéclauze, T.P. (2012). SPESS: A new instrument for measuring student perceptions in earth and ocean science. *Journal of Geoscience Education*, *60*(1), 83-91.

Lane, E., & Harris, S. (in prep). A new tool for measuring student behaviour engagement in large university classes.

Libarkin, J.C. & Anderson, S.W. (2005). Assessment of learning in entry-level geosciences courses: Results from the geosciences concept inventory. *Journal of Geoscience Education*, *53*(4), 394-401.

Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage Publications, Inc.

Marsh, P. (2007). What is known about student learning outcomes and how does it relate to the scholarship of teaching and learning? *International Journal for the Scholarship of Teaching and Learning, 1*(2), 1-13.

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

Merton, R.K., Fiske, P., & Kendall, P.L. (1990). *The Focused Interview: A Manual of Problems and Procedures*. New York: The Free Press.

Novak, G.M., Patterson, E.T., Gavrin, A.D., Christian W., & Forinash K. (1999). *Just in Time Teaching*. Upper Saddle River, NJ: Prentice Hall.

Ruiz-Primo, M.A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L.A. (2011). Impact of Undergraduate Science Course Innovations on Learning. *Science*, *331*, 1269-1270.

Simon B., & Taylor, J. (2009). What is the value of Course-Specific Learning Goals? *Journal of College Science Teaching*, November/December, 52-57.

Offerdahl, E. (under review). The process of science instructors' reflection: The role of a knowledgeable other.

Wieman, C. E., Perkins, K., & Harris, S. (2010). Transforming science education at large research universities: A case study in progress. *Change: The Magazine of Higher Learning*, *42*(2). http://www.changemag.org/Archives/Back%20Issues/March-April%202010/transforming-science-full.html

Wieman, C. E. (2009). Galvanizing Science Departments. *Science*, *325*(5945), 1181.

Wieman, C. E. (2007). Why not try a scientific approach to science education? *Change: The Magazine of Higher Learning*, *39*(5). http://www.changemag.org/Archives/Back%20Issues/September-October%202007/full-scientific-approach.html

Appendix 1

Results for the individual categories on the perception survey for the four geology courses. The results shown are % agree with "expert".

	GEOL 111	(2012)	GEOL 113	(2011)	GEOL 242	(2012)	GEOL 336	(2011)
	PRE (n=91)	*POST (n=53)*	*PRE (n=59)*	*POST (n=48)*	*PRE (n=85)*	*POST (n=62)*	*PRE (n=52)*	*POST (n=50)*
Overall	70.4	69.6	74.1	76.6	75.1	77.0	77.3	74.6
Memorisation	82.6	70.6	80.6	84.2	83.4	82.7	82.3	78.7
Science and Society	39.2	34.9	44.8	46.2	48.4	46.0	44.9	45.7
Mathematical Problem Solving	54.6	51.0	57.3	66.3	63.0	74.1	67.1	60.0
Personal Interest	89.8	90.8	93.1	93.3	93.6	94.8	96.2	93.8
Sceptical Reasoning	64.3	69.1	69.1	71.2	70.6	68.1	74.4	70.7
Conceptual Problem Solving	52.8	69.5	75.7	77.8	63.7	73.7	80.8	76.7
Human-Science Interaction	88.8	87.4	85.3	86.8	87.1	88.2	84.3	85.3

For more information go to www.akoaotearoa.ac.nz/projects/transforming-tertiary-science-education

AKO AOTEAROA
NATIONAL CENTRE FOR TERTIARY TEACHING EXCELLENCE

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

UC
UNIVERSITY OF CANTERBURY