

A methodology for improving active learning engineering courses with a large number of students and teachers through feedback gathering and iterative refinement

Iria Estévez-Ayres · Carlos Alario-Hoyos · Mar Pérez-Sanagustín · Abelardo Pardo · Raquel M. Crespo-García · Derick Leony · Hugo A. Parada G. · Carlos Delgado-Kloos

Accepted: 18 August 2014 / Published online: 25 September 2014
© Springer Science+Business Media Dordrecht 2014

Abstract In the last decade, engineering education has evolved in many ways to meet society demands. Universities offer more flexible curricula and put a lot of effort on the acquisition of professional engineering skills by the students. In many universities, the courses in the first years of different engineering degrees share program and objectives, having a large number of students and teachers. These common courses are expected to provide the students with meaningful learning experiences, which could be achieved by using active learning. The use of active learning in engineering courses improves traditional teaching by promoting students' participation and engagement, although active learning courses can be very sensitive to differences in learning paces or team conflicts; this being a challenge for the widespread adoption of active learning in courses with many students and teachers. This paper proposes a methodology that facilitates the detection and reaction to problems in active learning engineering courses with many students and teachers. This methodology is based on gathering feedback (from students and teachers)

This work has been funded by the Spanish Ministry of Economy and Competitiveness “EEE” project TIN2011-28308-C03-01, the Regional Government of Madrid “eMadrid” project S2009/TIC-1650, and the postdoctoral fellowship Alianza 4 Universidades.

I. Estévez-Ayres (✉) · C. Alario-Hoyos · R. M. Crespo-García · D. Leony · C. Delgado-Kloos
Department of Telematic Engineering, Universidad Carlos III de Madrid, Av. Universidad, 30,
28911 Leganés, Madrid, Spain
e-mail: ayres@it.uc3m.es

M. Pérez-Sanagustín
Department of Computer Science, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna
486, Edif. San Agustín, San Joaquín, Santiago de Chile, Chile

A. Pardo
School of Electrical and Information Engineering, University of Sydney, Sydney, NSW 2006,
Australia

H. A. Parada G.
Center for Open Middleware (COM), Universidad Politécnica de Madrid, Campus de Montegancedo,
28223 Pozuelo de Alarcón, Madrid, Spain

and decision-making processes at selected milestones. The methodology integrates intra-edition mechanisms in order to detect problems and react as the courses are being taught, and inter-edition mechanisms to ensure the persistence of necessary changes in the courses design. The methodology has been successfully applied during four consecutive editions to improve an undergraduate active learning programming course with an average of 257 students and 9 teachers per edition. An extended validation of expert educators suggests that this methodology can also be applied to traditional engineering courses.

Keywords Engineering course · Active learning · Feedback gathering · Iterative refinement · Large number of teachers and students

Introduction

At the present time, more and more universities are offering flexible curricula, allowing students to switch between degrees that belong to the same branch of knowledge during the first years (Stewart et al. 2013; Sursock and Smidt 2010). As a consequence, the first and even the second year of several degrees in the same branch share the same courses. These courses have many students enrolled and need to be taught simultaneously by many different teachers. One possible approach is to agree on the basic knowledge and skills to be acquired by the students and to give each teacher the freedom to choose specific contents, methodologies and criteria for passing a course. However, most universities are committed to ensure the same level of knowledge (and the same assessment system) for all the students enrolled in the same course, requiring a substantial coordination effort between teachers.

The former scenario is becoming more and more common in engineering schools, where the acquisition of general knowledge happens in the first years, while the specialization comes in higher courses (Alpay 2013). Moreover, in engineering degrees, the need for accompanying the lectures with applied knowledge, leads to the division of courses in theory and laboratory classes, usually taught by different people, adding an extra workload for the coordination between theory and laboratory classes within the same degree.

In addition to the general knowledge that students must acquire in these first- and second-year courses, students should develop other cross-curricular skills, such as leadership, initiative, critical thinking or teamwork, in order to become valuable professionals (Sheppard et al. 2008; Trilling and Fadel 2009; Litzinger et al. 2011). Introducing pedagogies, such as active learning (Felder and Brent 2009) and project-based learning (Walker and Leary 2009; Moursund 1999), in undergraduate engineering courses, can help students to acquire such skills. Active learning is useful to increase students' self-responsibility by promoting their participation during the course. Project-based learning is a particular active learning approach aimed at engaging students by means of realistic projects, similar to those found throughout their careers, that must be typically solved in small or medium-size teams. Both pedagogies facilitate the conceptual understanding of engineering principles and the development of cross-curricular skills, such as initiative or collaboration, and have been successfully applied in engineering courses (Martínez-Monés et al. 2005; Moura and Hattum-Janssen 2011).

Nevertheless, courses that use techniques from active learning are very sensitive to differences in students' backgrounds and learning paces, particularly when working in

teams. There are many issues (e.g., team conflicts, loss of interest, poor understanding of active learning...) that may arise during the course and the teacher should be aware of to react properly (Bouton and Garth 1983). Furthermore, if the course is taught by many different teachers, in different degrees, and includes lecture and laboratory groups (from now on, cohorts), then the coordination between teachers and awareness of what is currently happening in each cohort is of paramount importance. Therefore, a methodology suitable for active learning courses with many teachers and students that enables to detect these issues and adapt the course design to students' requirements and needs, even during its enactment, is needed.

This paper proposes such methodology, which allows the detection of problems and the dynamic adaptation of the main course settings (topics, learning activities, group configurations, assessment activities, etc.) as the course is being delivered. This methodology comprises intra- and inter-edition mechanisms. Both kinds of mechanisms are based on feedback gathering from students and educators and an analysis of this feedback. However, they differ on their goal: intra-edition mechanisms use this feedback to quickly react to problems during the course enactment, while inter-edition mechanisms analyze the changes and feedback obtained during the course enactment (i.e., the inputs and outputs of the intra-edition mechanisms), together with the feedback obtained at the end of the course, in order to ensure the persistence of the necessary changes in the course design. In order to be suitable for courses with many students and to enable a quick reaction from the teaching staff, the methodology relies on the usage of technologies for the automatic gathering of the feedback from students and the quick analysis of the feedback by the teaching staff. Changes in the course design are carried out as result of students' feedback, which is provided at predefined milestones. This way educators can react to reinforce those detected weaknesses and mediate to solve conflicts (e.g., those related to team work) before the course ends. Furthermore, the course design is refined every year to include students' feedback and educators' suggestions.

The methodology has already been applied to four editions of an undergraduate programming course that makes use of active learning and project-based learning pedagogies. On average on the 4 years under study, the course was by 9 teachers and had 257 students from 5 different degrees per year.

In this context the first research question is: *can the proposed methodology improve active learning engineering courses with a large number of students and teachers?* The answer to this question is researched employing data collected from four consecutive editions of the aforementioned undergraduate programming course. The second question addressed is: *can this methodology be applied to traditional engineering courses?* Forty educators with different degrees of expertise are consulted to assess the applicability of the proposed methodology to their own engineering courses.

The remaining of this paper proceeds with Sect. 2 summarizing the work related to course improvement through feedback gathering. Section 3 describes the overall methodology. Section 4 presents the undergraduate programming course employed as the case study, highlighting the changes introduced in its design during the 4-year span. Section 5 evaluates the methodology. Section 6 discusses the lessons learnt, and Sect. 7 draws the conclusions.

Literature on course improvement through feedback gathering

Currently, most universities collect feedback from students in an effort to monitor the quality of teaching. Actually, numerous universities under the umbrella of the EUA (the

European Universities Association, organization with more than 850 members from 42 countries) have joined efforts to describe and formalize quality assessment procedures (EUA 2009). However, the usage of Student Evaluation of Teaching (SET) to assess instructors or courses, i.e. for summative purposes, is, at least, a controversial practice (Balam and Shannon 2010): with numerous authors supporting the validity of such ratings (Cohen 1981), and detractors stating that this assessment is, at least, open to bias (Felton et al. 2004; Weinberg et al. 2007).

Despite their different opinions on the validity of SETs, most researchers consistently agree on the complex nature of the teaching activity, which has multiple and very different dimensions (e.g., organization, enthusiasm, teacher's personality); thus being difficult to capture this multidimensional nature using SETs (Marsh 2007). However, researchers also agree on the consistent correlation between SETs and other indicators of instructional effectiveness, such as students' scores, peer evaluation, etc. (Felder 1992), promoting the use of SETs as a reliable quantitative measure of teaching performance (Wachtel 1998).

In spite of the concerns raised by the usage of SETs for summative purposes, there is clear evidence that the use of student feedback for formative purposes can produce an improvement in the quality of teaching (Overall and Marsh 1979), this improvement being more significant when performed at midterm (Cohen 1980). Although faculty members report to care and to use SETs for improving teaching performance (Yao and Grady 2005; Ashton 2013), there is little advice in the literature on how to process feedback from students, except for reading it carefully and taking it seriously, a cumbersome task when performing multiple mid-term SETs in several large groups of students.

Some institutions with online courses, aware of the importance of both mid-term and end-term SETs, have developed systems for the online collection of student feedback (Bullock 2003), but the responses were handed over to the teachers with no categorization of the data or recommendations on how to analyze or to act upon the received feedback (Jara and Mellar 2010).

On the literature, instructors present some successful intra-edition experiences using mid-term SETs, e.g., the one carried by Steward et al. (2005) in an active learning engineering course. However, they do not propose a systematic methodology that can be applied to a large group of students. For example, in the aforementioned course, the number of students was between 14 and 25.

Brinko (1993), in her study about the effectiveness of the feedback collected from students to improve teaching, stresses that "the feedback is more effective when it is considered as a process, not a one-time quick fix" (suggesting the need for an iterative refinement), and that "feedback is more effective when is descriptive rather than evaluative". These findings are aligned with the usage of a continuous cycle of improvement through intra-edition mechanisms, such as the one proposed by Bateman and Roberts (1993), who combined fast-feedback and techniques from the area of management practices. These findings are also aligned with the usage of fast-feedback techniques such as mid-semester evaluation, informal early feedback or classroom assessment (Angelo and Cross 1993). The main idea of such techniques is to deploy a mechanism by which feedback (sometimes informal) is gathered from students with a frequency that allows instructors to reflect upon their answers and deploy, if needed, corrective measures. However, none of these techniques provide a formal procedure to process the feedback, once gathered from the students.

The authors of this paper proposed in Pardo et al. (2011) a fast-feedback mechanism that enabled teachers to categorize, and then analyze mid-term SETs to obtain qualitative information and deploy measures immediately. The methodology proposed in this paper is

an extension of that one, extending the feedback mechanisms, adding the inter-edition cycle and studying the improvement of a course after 4 years of enactment due to the use of this methodology.

The usage of inter-edition feedback mechanisms is more common and was subject of different proposals. For example, Takriff et al. (2011) reports an inter-edition methodology used in a Department of Chemical and Process Engineering aimed to continually improve the teaching and learning activities, and to ensure that the students achieve the intended learning outcomes in order to satisfy the accreditation requirements of their country. They only use student feedback obtained at the end of the semester, in the form of course assessments, student dialog sessions and an exit survey (performed by all the students of the final year). In that case, the improvement cycle involves faculty, students, an industrial advisory committee and external examiners. However, the definition of a mechanism for improving courses involving all the stakeholders of the University is out of the scope of this paper.

This work focuses on the internal mechanisms used by a department or by teachers within a course, as the one proposed by Barone and Lo Franco (2010). They proposed an inter-edition methodology, TESH (the Teaching Experiments and Student Feedback), which aims to continuously improve a course, but it is the teacher who decides the changes to be done, and then, measures the level of satisfaction of the students. The initial degree of satisfaction of the students is measured at the end of the first edition of the course, and then, the teacher may choose to alter aspects of the course through one or more teaching experiments (Barone and Lo Franco 2010) in the following editions.

To the best of our knowledge, there is no work in the literature proposing a methodology that includes inter- and intra- edition mechanisms to reflect upon the mid-term answers of students, their scores and end-term answers and deploy, if needed, corrective measures, suitable for courses with a large number of students enrolled and taught by many teachers.

A methodology for improving active learning engineering courses with a large number of students and teachers

The proposed methodology combines gathering feedback and decision-making processes to iteratively improve active learning courses where many teachers and students are involved. In such courses, several problems can arise (e.g. team conflicts, loss of interest, large differences between students' commitment and performance, etc.). These problems can be specific to one cohort, or be present at the same time in several cohorts, without the teaching staff being aware of them. Furthermore, the fact that there are different teachers can lead to unbalanced cohorts (e.g., due to different teaching styles). Usually at the end of the course, teachers only have a high level knowledge of the problems of her/his group/s of students, but they do not know either the experience of the rest of the teaching staff, or the experience of the students (Brookfield 1995).

As shown in Fig. 1, this methodology overcomes the aforementioned limitations in active learning courses with many students and teachers by integrating intra-edition mechanisms in order to detect problems and react during the course enactment, and inter-edition mechanisms to ensure the persistence of necessary changes in the course design.

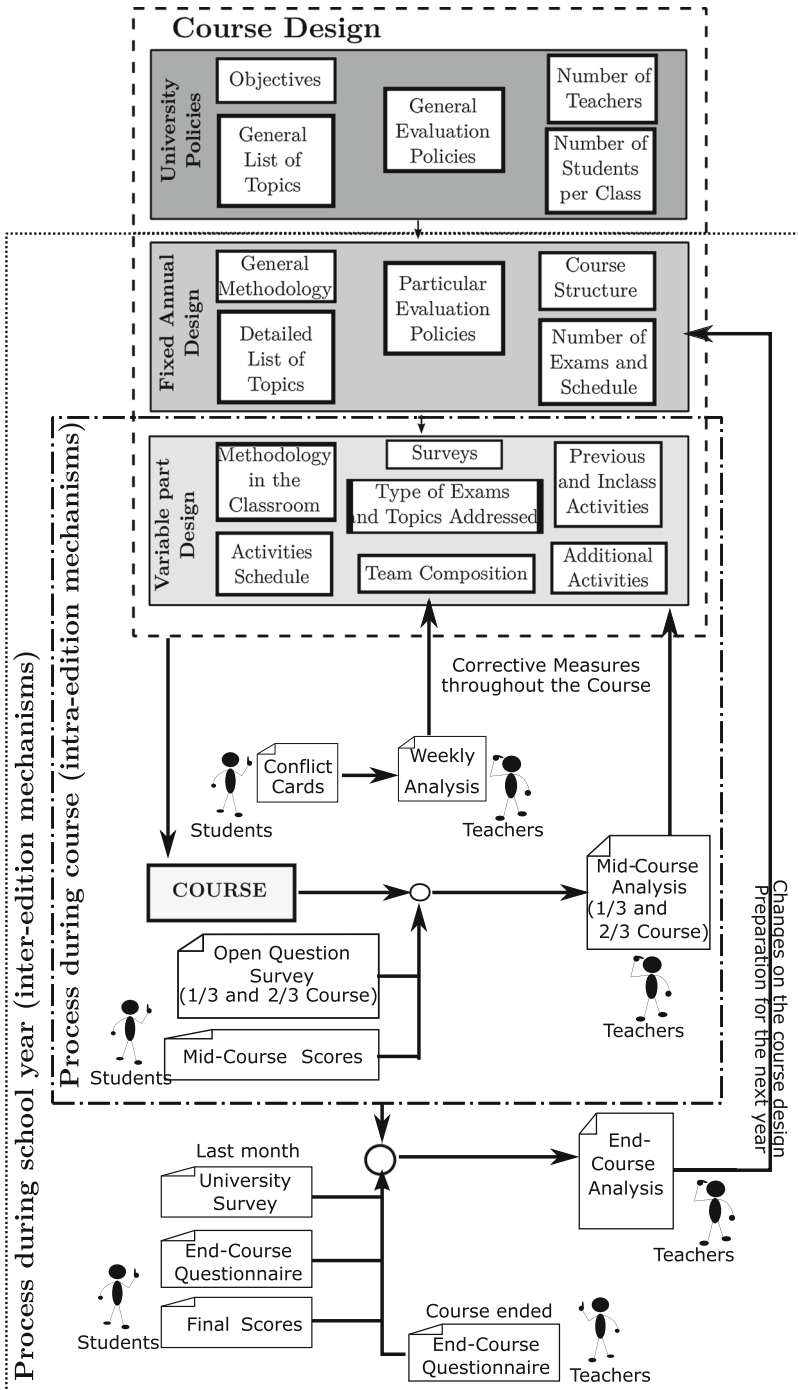


Fig. 1 Overall methodology and impact on the course design: process during course (intra-edition mechanism) and process during school year (inter-edition mechanism)

Intra-edition mechanisms

Intra-edition mechanisms are introduced since the beginning of the course, so that instructors are aware of the main issues that arise during the course. The different sources that provide data are: *students' opinions*, *mid-course scores* and *team conflict cards*.

Students' opinions are collected using voluntary, anonymous, web-based open question surveys. These surveys are posed twice during the course (at one third and two thirds of its duration) and the students are asked only two questions: "Tell us the most positive aspect of the course (since the last survey)" and "Tell us the most negative aspect of the course (since the last survey)". This way, students should reflect on all the aspects of the course before answering these questions, affecting positively the students' engagement (Daly 2008).

Students' mid-course scores are gathered from weekly assessments. Every summative assessment should be taken into account, including mid-term exams, laboratory assignments, and other sources of assessment (if the teaching staff decides so).

Students' opinions and scores are analyzed throughout the course. Teachers meet at least twice during the course (when the students' opinions from the different open question surveys are available) to identify assets and pitfalls. Conclusions of the analysis are discussed with students, and appropriate reaction mechanisms are immediately applied, depending on the specific demands (e.g., solve more exercises in class, increase the duration of the exams, organize reinforcement classes...).

In addition, if the course includes the development of complex team projects, which are quite common in engineering courses, then teachers can gather more feedback from conflict cards. *Team conflict cards* are based on the classification defined in Oakley et al. (2004), where four different problems within a team of students were identified: presence of hitchhikers (students that refuse to do their share of work); domineering team members who try to do everything their way; resistant team members who resent having to work in a team and try to sabotage the team effort; and team members with widely divergent goals.

The conflict cards are weekly collected during the project development, and used to monitor the evolution of the teams work. A conflict card is handed over to each student with the number of her/his team and the four aforementioned problems with a Likert-4 scale. Each student should state the level of each the problem from 0 (no existence of this specific problem) to 3 (this problem is really jeopardizing the team performance). These conflict cards are weekly analyzed to enable teachers to react immediately to conflicts within teams. Also, conflict cards allow evaluating whether the distribution of teams is appropriate. This mechanism can be useful in active learning courses where students should solve a realistic project in teams (project-based learning). In this kind of courses, team conflicts can arise and the teaching staff is not usually aware of them, and the students do not reflect on their own behavior. The conflict cards serve for a twofold purpose: encouraging the communication students-teacher; and, making each student aware of her/his team dynamics and of her/his own performance within the team.

Inter-edition mechanisms

Apart from quick reaction to problems arisen during the course, the proposed methodology aims to provide long-term solutions too. Each course edition must be thoroughly analyzed in order to refine its design for future years. This refinement is built up on the results of intra-edition mechanisms together with the following information gathered at the end of

the course and not taken into account by the intra-edition mechanisms: *end-course students' opinions*, *teachers' opinions* and *students' final scores*.

End-course students' opinions are collected both from traditional university surveys and an anonymous, voluntary, web-based end-course questionnaire.

Usually, traditional university surveys include Likert-5 scales about their level of satisfaction with the teaching staff, the evaluation policies, the course workload and the acquired level of different skills, and open text questions (e.g., "Tell us what you would improve in this course" or "Tell us what you would maintain in the course"). Other sources of feedback coming from the University and different from surveys can be added to this methodology as inter-edition mechanisms, depending on particular institutional contexts.

The end-course questionnaire includes several open questions to collect the most positive and most negative aspects of the course regarding the teaching staff, the collaborative learning experience (if any), and students' general opinion regarding the whole course; and Likert-5 scales for any other aspect of interest. These scales can change depending on the edition of the course. If it is the first edition, maybe teachers want to cover more aspects, from pedagogical aspects ("Indicate the level of usefulness of the previous activities?") to technical aspects ("Did you find useful the course virtual machine?"), while in other editions the staff may want to focus on new additions to the course ("Did you find the lab exam rehearsal useful?") or even asking specifically to students that are retaking to compare issues from different editions ("What virtual machine desktop do you prefer, Kde or Gnome?"; "Do you feel that the workload increased, decreased or was equal as compared to the previous year?").

Teachers' opinions are gathered from an end-course questionnaire where they propose 3 issues to maintain and 3 issues to change for the next edition, and a Likert-5 survey regarding the level of satisfaction with the enactment of the course, tools employed and overall course design.

Students' final scores are also collected with a complete record of students' grades throughout the course. Every summative assessment with its weight should be taken into account: individual tests, group submissions, lab exams, etc.

Once all the data are available, the action plan consists of:

1. *Analysis and discussion*. The teaching staff jointly analyzes and discusses all the data sources in a first meeting 1 month after the course. In this meeting, the scope of the changes is narrowed. If needed, a meeting with the student representatives is carried out.
2. *Decision-making*. In a follow-up meeting, major decisions about the course design are taken, such as schedule changes, evaluation policies or general contents. Students are informed about these decisions through the University.
3. *Application of changes*. Finally, several extra meetings are scheduled to apply changes in the design, ensure the improvement of materials and their fitting to the new adjustments made to the course.

Case study

This methodology was applied to four consecutive editions (2009 to 2012) of a semester programming course, taught in five different degrees in Telecommunication Engineering by several teachers with different profiles. On average, 257 students enrolled per year (see

Table 1). Following university policies, students were organized in large and small groups according to the degree they enrolled. Lectures were delivered in large groups and had a limit of 120 students, with 5 large groups, one per degree. Students attended laboratory lessons in smaller groups (up to 40 students). On average, the students were divided into 9 laboratory groups. The same university policies state that only teachers holding a PhD could teach large groups, while in lab groups there was no restriction. So, the teaching staff had also different profiles: a senior lecturer, several junior lecturers (5), teaching assistants (2) and part-time lecturers (3). On average, 9 teachers were involved per course. The course had to apply a continuous assessment scheme and provide the following learning outcomes:

1. design and development of applications in C programming language;
2. use of tools for proficient application development (compilers, debuggers, IDEs, etc);
3. employment of teamwork techniques to develop an application for mobile devices;
4. use of self-learning techniques.

The course design followed an active learning approach: students were required to work at home on several activities prior to face-to-face sessions. These activities introduced them to the topics of the current week, covering theoretical concepts. Students were expected to solve questions by themselves or asking the instructor in a tutoring session or in the online course forum. In face-to-face sessions the instructor assumed that theoretical concepts had been already covered by students; so, there was no theoretical explanation unless students requested it. Sessions were mainly dedicated to solve programming problems. During the first half of the course, students worked in pairs in laboratory assignments. For the second half, and in order to foster teamwork, instructors rearranged students in teams of four to carry out a realistic programming project. The continuous evaluation consisted of 8 team tests and 9 individual tests. The tests were both practical and theoretical. The theoretical tests were mainly individual and problem-based; while the practical tests comprised two group project submissions, an individual project test, submissions of code in pairs and a presentation of the work at the end of the course by the whole group. All these tests were taken into account to calculate the final scores.

This general course design was refined during 4 years applying the aforementioned methodology. Table 2 shows the adjustments in the course design as part of its continuous

Table 1 Number of different teachers, groups (theoretical and laboratory), enrolled students, students that started the project and students that passed the course by edition

Ed.	Num teachers	Groups		Number of students				
		Th	Lab	Enrolled	Started the project		Passed the course	
					N	%	N	%
2009	7	5	8	198	152	76.77	122	61.62
							84	42.42
2010	8	5	8	248	205	82.66	123	49.60
2011	10	5	11	322	273	84.78	147	45.65
2012	6	5	9	260	202	77.69	120	46.15

In italic, the data of 2009 if the threshold had been applied

Table 2 Problems and main adjustments in the design of the course during the four editions (2009, 2010, 2011 and 2012), and mechanisms that helped detect the problems

Problem	Mechanism		Adjustment (and edition)
	Intra	Inter	
(a) Students neither like nor understand active learning	✓		(a) Devote time in class to explain work habits required by the course (2009). (b) Explain work habits in the first session; brief reminders during the course (2010)
(a) Students are hesitant to ask questions during lectures. (b) Complaints about having unresolved doubts	✓	✓	Start theoretical sessions with a brief explanation of the needed concepts and ask for doubts to the students (2010)
(a) Students do not prepare the required exercises before the class. (b) Complaints about not providing enough solved exercises	✓	✓	(a) Use of forums to solve exercises following an active learning approach: teachers propose exercises and students must solve them prior to the revision by the teacher (2010). (b) Add self-evaluation questions in the web material (2011)
Students do not attend to individual tutoring (only when close to the end of the course or the project deadline)		✓	Additional weekly group tutoring sessions (50 – 60 % of the students attend to these sessions) (2011)
Conflicts within teams (some students take advantage of their team mates)	✓	✓	(a) Establish an individual score threshold to pass the course (2010–2012). (b) Individual project exam (2010). (c) Change the policy on how teams are composed (students with similar achievements are grouped together) (2010). (d) Usage of conflict cards to monitor problems within teams (2010). (d) Define an 'Internal Teachers' Protocol on how to deal with team conflicts (2011)
(a) Different paces and learning styles. (b) Beginning the evaluation so early in the course turns out to be a disadvantage for students		✓	Adjustments on the distribution of grades during the course: shift most of the weight to the final part (in 2012, 63 % of the final score was obtained in the last month), while maintaining the motivation to study with several assessments (with less weight) during the rest of the course (2010–2012)
High course workload	✓	✓	Reduce the number of assessment activities (from 17 to 14) (2010, 2011)
(a) Complaints about the project rubric being too generic. (b) Rubric gives excessive degrees of freedom. (c) Students do not understand their errors. (d) Students massively ask for a review of the project grade	✓	✓	(a) Revision of the project rubric (2010). (b) Creation of particular rubrics for each assessment activity (2010). (c) Publication of all the evaluation rubrics (2010). (d) Formative feedback to the students about their errors (2011)
Complaints about lack of information on the individual project test	✓		Individual project test rehearsal (voluntary, without score) (2012)

Table 2 continued

Problem	Mechanism		Adjustment (and edition)
	Intra	Inter	
Concepts explained at the beginning of the course are easily forgotten; students fail to apply them to the project	✓		(a) Disassociate large and small group sessions (2010). (b) Review of exercises that they should apply to the project during the large group sessions (2011, 2012)
(a) Students are not familiar with Linux. (b) Students used to work with an IDE (Eclipse) in earlier courses (thus they are not familiar with command-line)	✓		(a) Insertion of a Linux Introduction Session at the very beginning of the course (2010). (b) Use only Linux command-line instead of an IDE, i.e. use <code>emacs</code> , <code>gcc</code> , <code>gdb</code> , <code>svn</code> instead of Eclipse built-in tools (2010)
(a) The course schedule delays programming the first 3 weeks. (b) Scores show that students need more programming	✓		Mix explanations of supporting tools and actual programming from the very beginning of the course (2012)
Information gathered by the surveys is incomplete	✓		(a) Inclusion of more Likert-5 questions (2010). (b) Focus on identified problems of the current course (2012)

improvement and the kind of mechanism that helped detect the existing problems and take actions on them.

For example, during the first edition, it was detected that some students took advantage of their team mates (see Table 2): several teams complained in the questionnaires and personally to the teaching staff about their colleagues (their lack of commitment, having different objectives, etc.). After an analysis of the students' final scores, teachers found that several students passed the course with very low scores in the individual tests (and possibly with a low knowledge of the course contents) but high scores in the team project.

So, different measures were taken:

- It was noticed that teams complained late: only when the project deadline was close and little could be done by the teaching staff. So, it was decided to encourage the communication students-teacher introducing the *weekly Conflict Cards*, which were not present in the first edition of the course.
- During the first year of enactment, two different policies were used to form teams: students with similar achievements grouped together; and mixing students with different achievements. It was detected that a higher percentage of teams did not agree with the second option, and, also that the level of students' satisfaction with their team colleagues was lower with the second option. So, it was decided to *implement only the first criterion* in the following editions.
- In order to avoid that students took advantage of their team mates, *an individual project exam and an individual score threshold to pass the course (50% of the individual points) were established*. This last decision explains the decrease on the percentage of students passing the course shown in Table 1 (from 61.62% in 2009 to 49.60% in 2010). However, if the individual threshold had been applied to the scores in the first year, the number of students that would have passed the course would have decreased from 61.62% to 42.42% (as shown in Table 1).

Evaluation

The evaluation of the methodology is organized into two sections, each focused on one of the two research questions of this paper. Since the objective is to analyze the impact of the proposed methodology in authentic courses including many factors and contextual issues, we compared and triangulated the data extracted from the different information sources using the mixed method proposed in Martínez-Monés et al. (2006). These information sources contained qualitative data, which helped to identify tendencies of the intervention in this case study drawing on the strengths and weaknesses (Gahan and Hannibal 1999; Denzin and Lincoln 2005), and quantitative data, which served to reinforce or discard each of the detected tendencies. Each section details the qualitative and quantitative data employed to extract conclusions according to the nature of the research question addressed.

Evaluation of the methodology in the case study

This section evaluates whether the methodology proposed improved the active learning programming course used as the case study. Students' answers to surveys and questionnaires during the four course editions support this evaluation (see Table 3). The most

Table 3 Students' surveys and questionnaires used in the evaluation (where XX is the year of enactment: 2009, 2010, 2011, 2012)

	When?	What?	By Whom?
CF-st-XX-1	1/3 of the course	Open question survey	Teaching staff
CF-st-XX-2	2/3 of the course	Open question survey	Teaching staff
UF-st-XX	4/5 of the course	Likert scales + Open questions	University
CF-st-XX-3	End of the course	Likert scales + Open questions	Teaching staff

positive and negative aspects according to students are employed as qualitative data. The proportion of positive and negative comments related to these aspects are used as quantitative data (see Table 4). Three different researchers participated in the data analysis and in the extraction of findings.

Notice that all the students' opinions (both the open questionnaires used in the intra-edition mechanism and the surveys used in the inter-edition mechanism) are voluntary, anonymous and web-based, i.e. even the university surveys are willingly accessed and filled in by the students. Usually, students fill them in from their homes. Each questionnaire and survey informs the students on how their data will be used. Concretely, the following message is shown into the web survey: "This survey is voluntary and anonymous. The data collected will not be used as part of the assessment of the course, but only to support the learning process. Furthermore, after an aggregation process, the data will be used for research about future improvements in the methodology and contents of this and other similar courses".

Given the optional nature of the university surveys, the average percentage of students that answer them is low, even though the University encourages and reminds them. Students typically do not fill the surveys, unless they are really pleased or really upset by the course or the staff. Given this reality, the University established in an internal regulation of 2012 that surveys filled in by > 15 % of students, with a minimum of 5, will not be considered representative and the teaching staff will not know the results. Surprisingly, as shown in Table 4, in all the editions of the course, we found that more students answered the specific surveys from the course than the generic surveys from the university.

The first step of the analysis was the definition of two *information questions* (IQ) (Denzin and Lincoln 2005) derived from the first research question, to help drive the data comparison: (IQ1) *Are the intra-edition mechanisms a good technique to improve the course during its enactment?* and (IQ2) *Are the inter-edition mechanisms a good technique to ensure the persistence of necessary changes in the course design?* Each question was used to create a set of categories that facilitated the classification of data related to the course: general methodology, theoretical sessions, lab sessions, previous and additional activities, course changes, teaching support, collaborative learning, workload and evaluation. The NVivo software (Gahan and Hannibal 1999) was used to classify all the data from the four course editions according to these categories. The second step was to structure the data processed with NVivo in two different tables, each corresponding to one information question. For the first information question, data from the surveys of the same course editions were compared. For the second information question, the data compared belonged to different editions. This organization enabled two of the researchers to derive their own set of partial results. Finally, in the last step, these two researchers discussed their partial results with a third researcher and they jointly extracted the final findings (see Tables 5 and 6).

Partial results related to IQ1 (1.1–1.3 in Table 5) indicate that the intra-edition mechanisms were effective for identifying the most problematic aspects of the course concerning the general methodology, lab sessions or teaching support, and react consequently to address them during the enactment. First, partial result 1.1 indicates that the methodology employed enabled to get an overview of students' satisfaction at different moments during the course enactment. Quantitative data supporting this result point out that the number of complaints decreased through the course in most editions. For example, in the 2009 edition, 48 % of the comments were complaints at the beginning with only 32 % at the end (Table 4). This decrease is not observed in the 2012 edition, since the number of complaints from the beginning of the course increased at the end. However, in this edition, we could appreciate that most students agreed with the course design and enactment after comparing the overall percentage of students' positive comments (75 %) and negative ones (25 %) (see Table 4). Second, partial result 1.2 denotes that the methodology employed enabled, not only to detect the percentage of complaints and successful aspects, but also to identify the most relevant issues that students asked for improvement. For example, looking at students' comments, we could observe that one of the main aspects for them was the high workload, which was higher than in most undergraduate courses due to the use of active learning. Nevertheless, teachers took actions to gradually decrease the number of assessment activities (from 17 to 14), and the students related that with a lower average workload, as it can be seen with the reduction of complaints concerning this issue in Table 4. Finally, data supporting partial result 1.3 suggests that the intra-edition mechanisms were a good approach to quickly react to students' needs while running the course (see for instance in 2009: "*It has been one of the few times that students' complaints were taken into account for the enactment of the course, if not in everything, in some aspects, like the workload.*" [CF-st-2009-2]).

Partial results related to IQ2 (2.1–2.4 in Table 6) show that the inter-edition mechanisms were a useful approach to identify what successful aspects had to be maintained from one edition to another and what aspects needed to be revised for improvement. First, quantitative data supporting partial result 2.1 indicate that the difference between the percentage of positive and negative aspects highlighted by the students during the 4 editions grew with time (60 % of positive aspects in 2009, 63 % in 2010, 65 % in 2011 and 75 % in 2012). These results suggest that the overall students' satisfaction was improved through the different editions. Second, after a deep analysis of students' comments (see selected data related with partial result 1.3 in Table 5), partial result 2.2 points out that the nature of complaints was different from one course to another and that some of the problems were solved. While in the first edition students complained about the extension of the curriculum, the difficulty of the project and the amount of exams, these topics were not repeated in any other edition. Nevertheless, the disappearance of complaints about most aspects had a side effect in subsequent editions: most complaints grouped around one single aspect, the evaluation, which is partially constrained by university policies. Therefore, despite teachers' efforts to reduce the number of tests, students kept complaining more and more about this aspect (Table 4). Third, we also observed that students from editions 2011 and 2012 explicitly expressed their satisfaction regarding the overall course methodology and organization (partial result 2.3). For instance, one student reported: "*The course planning and enactment were in general very successful.*" [CF-st-2012-3]. Finally, selected comments supporting partial result 2.4 evidenced that the improvement from one edition to another was noticeable by those students enrolled in the course for a second time: "*Theoretical sessions this year improved with respect to previous editions*" [CF-st-2012-2].

Table 4 Students' comments from 2009 to 2012 (quantitative data)

Aggregated comments excluding workload and evaluation. Nodes in NVivo project: general methodology, theoretical sessions, lab sessions, previous and additional activities, course changes, teaching support, collaborative learning

Ed.	CF-st-1 <i>T: (P, N)</i>	CF-st-2 <i>T: (P, N)</i>	UF-st <i>T: (P, N)</i>	CF-st-3 <i>T: (P, N)</i>	Total <i>(P, N)</i>
09	110: (52, 48)	65: (68, 32)	52: (42, 58)	52: (81, 19)	(60, 40)
10	104: (64, 36)	55: (78, 22)	64: (55, 45)	28: (75, 25)	(63, 37)
11	155: (69, 31)	56: (80, 20)	67: (58, 42)	82: (72, 28)	(67, 33)
12	150: (81, 19)	32: (81, 19)	NA	79: (61, 39)	(75, 25)

Comments related to workload (all the comments are negative). Nodes in NVivo project: workload

Ed.	CF-st-1 <i>T</i>	CF-st-2 <i>T</i>	UF-st <i>T</i>	CF-st-3 <i>T</i>	Total <i>T</i>
09	77	23	34	14	148
10	77	24	31	17	149
11	71	33	24	11	139
12	68	14	NA	7	89

Comments related to evaluation. Nodes in NVivo project: evaluation

Ed.	CF-st-1 <i>T: (P, N)</i>	CF-st-2 <i>T: (P, N)</i>	UF-st <i>T: (P, N)</i>	CF-st-3 <i>T: (P, N)</i>	Total <i>(P, N)</i>
09	23: (52, 48)	7: (57, 43)	9: (44, 66)	4: (25, 75)	(49, 51)
10	9: (11, 89)	6: (50, 50)	26: (11, 89)	11: (9, 91)	(15, 85)
11	19: (47, 53)	12: (17, 83)	34: (6, 94)	7: (14, 86)	(19, 81)
12	24: (17, 83)	8: (0, 100)	NA	9: (0, 100)	(10, 90)

For each cell, $T: (P, N)$ where T is the total number of comments; P , the % of positive comments and N , the % of negative ones. Note that the workload only received complaints and that data from UF-st-2012 was not available

Evaluation of the usefulness of the methodology in traditional engineering courses

The proposed methodology and the adjustments in the design of the case study after the four editions were evaluated by peers: expert educators in engineering courses. The evaluation consisted of a survey with 6-point Likert scales with selected assertions and open text questions for further clarifications. The evaluation was voluntary, web-based and anonymous, and was distributed among the teaching staff of six schools of engineering from six different Spanish universities.

The 40 educators that answered the evaluation had a different degree of expertise (78 % of them with more than 6 years of teaching experience), and were used to teach a large number of students (more than 40 for 88 % and more than 160 for 30 % of the teachers

Table 5 Final findings and partial results of the qualitative analysis for IQ1

Findings	Partial results	Selected data
<p><i>IQ1 - Intra-edition mechanisms were a good technique to improve the course during its enactment, since in the case study they enabled to:</i> (a) identify students' complaints throughout the course about the general methodology; (b) react quickly, taking decisions to decrease the number of complaints about the general methodology; (c) detect what aspects were relevant for students and keep an eye on them during the rest of the course; (d) increase students' perception that teachers take into account their opinions; (e) determine the course organizational problems; (f) find out the evaluation aspects that had to be improved</p>	<p>1.1. The number of students' complaints about the methodology decreased throughout the course in 2009, 2010 and 2011 editions</p> <p>1.2. Students' comments in the different editions were of different nature and were addressed from 1 year to another. Main complaints by number of comments per year were: - 2009 (1) too extensive curriculum; (2) workload; (3) few number of solved problems; (4) lack of teacher support in theoretical classes; (5) the difficulty and overload of the project; (6) team work conflicts; (7) scoring system; and (8) number of exams. (9) 2010 (1) concrete course topics; (2) workload; (3) few number of solved problems; (4) team work conflicts; and (5) lack of teacher support in theoretical classes. - 2011 (1) workload; (2) few number of solved problems; and (3) team formation policies. - 2012 (1) workload; (2) few number of solved problems; (3) team formation policies; and (4) concrete course topics</p> <p>1.3. Students reported that their opinions about the course had been taken into consideration for changing some course aspects</p>	<p>See quantitative results in Table 4</p> <p>"Too much syllabus in so little time and very few programming practices." [CF-st-2009-1] "This course, unless you understand it well, is very hard; and it can be super difficult to understand. Therefore, it is necessary to explain things more carefully, not just that we read the previous activities and try to do the exercises, because, as usual, we will not know how to do them." [CF-st-2009-1] "There are not enough explanations in the classroom. We do not have the solutions to problems at hand, and in the forum all the doubts are not resolved." [CF-st-2010-1] "Some topics necessary to advance the project, such as <code>hread</code> or <code>fwrite</code>, were explained in theoretical lessons." [CF-st-2010-3] "Uneven distribution of groups." [CF-st-2011-2] "Previous activities are effective." [CF-st-2011-1] "The workload has gradually decreased." [CF-st-2011-2] "In contrast to the good content, the downside is that some of the previous work is too long, but I really think that is just the other side of the coin, you cannot have such good activities without them being long." [CF-st-2012-1] "There is a very good communication between students and teacher in the forum and in the lab sessions. The course is very dynamic, so it evolves in a good way." [CF-st-2012-1]</p> <p>"It has been one of the few times that students' complaints were taken into account for the enactment of the course, if not in everything, in some aspects, like the workload." [CF-st-2009-2] "I noticed changes for the better in the evaluation, not so many exams at the end of the class. Besides, previous activities forced us to study the theory and that is a weight off." [CF-st-2009-2] "The workload has gradually decreased." [CF-st-2011-2]</p>

asked), thus demanding a proper methodology to cope with issues that typically arise in these courses, such as differences in learning paces and conflicts when working in teams.

For example, the use of open question surveys filled out by students at selected milestones during the course in order to detect problems was positively assessed by 95 % of the surveyed educators, with 78 % of them finding positive the employment of conflict cards. Further, 95 % of the answers were in agreement or complete agreement that the teaching staff should periodically reflect about the achievement of the established course objectives; the inter-edition meetings designed for this purpose received 98 % of positive critiques. Explicit comments made by these educators confirmed the interest on this methodology, although there were some doubts about the workload it entails, especially when few teachers are in charge of a course with many students: *“I think this methodology is very interesting from a teacher’s perspective; however, it imposes an overload hard to assume in situations with few teachers”*; *“I consider this is an excellent methodology but requires a high dedication from the teaching staff”*.

The current course design, product of 4 years iterating on this methodology, was presented to the 40 educators, obtaining also positive reviews. For instance, 90 % of the educators thought that a similar design could be useful, in general, in their courses (17.5 % completely agree, 52.5 % agree and 20 % somewhat agree). Also, 85 % of them were positive about the opportunity this design offers to increase students’ engagement. As a counterpart, explicit comments made by the educators pointed out the high time and effort it may take to enact this course, particularly regarding students’ assessment, e.g. *“the assessment load is somewhat high”*; *“it seems a good approach, but I am not sure about its sustainability from the teaching load perspective”*. Everything considered, experts’ comments praised the current course design in general, although some refinement is still needed to adapt it to different time constraints.

Discussion

The evaluation results from the previous section allow us to answer the research questions defined at the beginning of this paper.

Regarding the first research question *can the proposed methodology improve active learning engineering courses with a large number of students and teachers?*, the proposed methodology was found to help improve the active learning engineering course employed as a case study here over a period of 4 years:

- The fact that this course was mandatory, had a large number of students, and was simultaneously delivered in several cohorts from different engineering degrees made it representative to determine the extent to which this methodology can be useful.
- This course happened to be designed from scratch in its first edition, and 4 years later the concerns raised by the students, identified thanks to this methodology, had been addressed and solved (see Table 2).
- These adjustments enabled that, after the 4 years of this course, students almost stopped complaining about active learning, the course organization or the teaching support, and focused their negative comments on one single aspect, the evaluation (see Tables 4 and 5).

Regarding the second research question *can this methodology be applied to traditional engineering courses?*, the main findings after the peer evaluation of the methodology by 40 engineering educators (see Sect. 5.2) indicate that:

Table 6 Final findings and partial results of the qualitative analysis for IQ2

Findings	Partial results	Selected data
<p><i>IQ2 - Inter-edition mechanisms were a good technique to ensure the persistence of necessary changes in the course design, since in the case study they allowed to strengthen changes related to: (a) the methodological and organizational problems; (b) the evaluation issues; (c) other topics and aspects that required a special attention, such as materials, team work conflicts, etc</i></p>	<p>2.1. The difference between the percentage of positive and negative comments increased with the editions, 2012 being the one with a higher percentage of students' positive comments</p> <p>2.2. Students' comments between editions were of different nature and were addressed from 1 year to another</p> <p>2.3. Students' comments in 2011 and 2012 editions show that they liked the overall course methodology</p> <p>2.4. Students repeating the course for the second time appreciated the changes from one edition to another</p>	<p>See quantitative results in Table 4</p> <p>See selected data supporting partial result 1.2 in Table 5</p> <p><i>"The course planning and enactment were in general very successful."</i> [CF-st-2012-3] <i>"Entertaining and well-explained lectures, active and quick forum, that is, if you write a doubt, you get immediately an answer. This way, you can follow the course day-to-day."</i> [CF-st-2011-1]</p> <p><i>"Theoretical sessions this year improved with respect to previous editions."</i> [CF-st-2012-2] <i>"This year, the explanation of the project improved with respect to the previous edition. I understand better what I should do."</i> [CF-st-2012-2]</p>

- The methodology received positive evaluations from the vast majority of the experts.
- The consulted educators use different methodologies. Nevertheless, most of them (90 %) found that the methodology proposed in this paper could be useful in their own courses.

One of the main criticisms received from the expert evaluation was the additional workload of this methodology on students and teachers. This additional workload will strongly depend on the maturity of the course. For example, during the first editions of a course, the structure and most materials need to be developed from scratch. These contents need to be refined in the following editions until reaching a state of maturity. Consequently, if teachers apply this methodology, a large number of complaints from the students will be related to this issue (e.g. complaints about the extension of the curriculum, number of exams, quality of the materials, etc.) and the teachers will be able to react and improve the different aspects even during the course enactment. This additional workload during the first editions of the course, facilitates reaching a state of maturity quickly, and reduces the time to provide materials with a high quality. When the course has reached a state of maturity, the workload of processing students' opinions is much lower. Moreover, as the course evolves and the number of changes between editions decreases, the workload can be reduced by gradually lowering the frequency of the intra-edition surveys. The same happens with the number of aspects that students must assess in the end-course questionnaire. For example, those aspects not mentioned by the students in the open question surveys delivered during the course could be avoided, focusing on the more problematic issues raised on the surveys or on their opinion about the changes introduced in the current edition.

When asked about the evolution and improvement of a course, students repeating the course for the second time possess a wider perspective and their informed opinions are very valuable. In order to reduce the actual number of questionnaires to process, one possible approach is focusing on gathering feedback from repeaters and also maintaining a reduced control group of students taking the course for the first time.

Finally, team conflicts usually arise within the first (forming) and second (storming) phases of group work identified by Tuckman (1965). So, teams should be more carefully monitored for team conflicts during these two phases. From the experience through these four editions, it is very difficult to foresee the exact time when these two phases happen, as the internal pace of each team is different. However, if the course has intermediate deliverables, we observed that the conflicts arise around the deadline of the first submission and the teams that overcome these problems quickly evolve to the third (norming) and fourth (performing) phases of Tuckman model. So, it is our belief, that teachers can reduce the number of weekly conflict cards to process using them until students submit the first team work.

When applying this methodology it is also crucial that teachers have both the will and the opportunity to make changes during the course in order for the methodology to be of value. For example, some institutions may hinder, and even prevent, any changes in the course during its enactment. So, it is desirable to have institutional support or at least institutional flexibility regarding the improvement of a course while it is being taught.

Everything considered, we, as teachers conducting the course in which this methodology was applied, believe that the additional workload is worthy. We believe that it helps detect the wide variety of problems that happen in an active learning course in which many students and teachers are involved, providing also a more concrete vision of the reality in the classroom. Moreover, actively involving students in the course design process by using

this methodology, we found that not only improves the teaching of the course, but also their engagement and commitment (see finding 1.3 in Table 5), finding that is aligned with the results from Prince (2004).

Conclusions and future work

This paper presented a methodology based on gathering feedback and iterative refinement that enabled to improve an active learning programming course in which many students and teachers were involved. This methodology led to the detection of problems of different nature during four consecutive editions, such as team working problems, methodological and organizational problems and evaluation issues. The vast majority of these problems were solved and the teachers are still working on the improvement of the course. The complementary evaluation performed by expert educators suggests that this methodology can also be employed in traditional engineering courses with similar characteristics.

As part of the future work, this methodology will be implemented in traditional engineering courses with the support of some of the experts who carried out the peer review, in order to note differences with the case study. Also, an ongoing research line is working on facilitating teachers and students the visualization of the aggregated data, arranged according to the categories defined for the qualitative analysis, making use of the visualization tool proposed in Leony et al. (2012).

Acknowledgments The authors want to thank the teachers involved in the 4-year enactment of this course. We appreciate all your time and effort. We also want to thank to the external teachers that completed the survey. Thank you for taking the time to help us improve our teaching.

References

- Alpay, E. (2013). Student attraction to engineering through flexibility and breadth in the curriculum. *European Journal of Engineering Education*, 38(1), 58–69.
- Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers*. New York: Jossey-Bass.
- Ashton, K. B. (2013). Reflections on the gathering of student feedback: An evaluative case study. *Higher Education Research Network Journal*, 7, 3–12.
- Balam, E. M., & Shannon, D. M. (2010). Student ratings of college teaching: A comparison of faculty and their students. *Assessment & Evaluation in Higher Education*, 35(2), 209–221.
- Barone, S., & Lo Franco, E. (2010). TEF methodology for statistics education improvement. *Journal of Statistics Education*, 18(3), 1–25.
- Bateman, G.R., Roberts, H.V. (1993). TQM for professors and students.
- Bouton, C., & Garth, R. Y. (1983). Students in learning groups: Active learning through conversation. *New Directions for Teaching and Learning*, 1983(14), 73–82.
- Brinko, K. T. (1993). The practice of giving feedback to improve teaching: What is effective? *Journal of Higher Education*, 64(5), 574–593.
- Brookfield, S. (1995). The getting of wisdom: What critically reflective teaching is and why it's important. In: *Becoming a Critically Reflective Teacher*, Jossey-Bass.
- Bullock, C. D. (2003). Online collection of midterm student feedback. *New Directions for Teaching and Learning*, 2003(96), 95–102.
- Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. *Research in Higher Education*, 13(4), 321–341.
- Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research*, 51(3), 281–309.
- Daly, C. (2008). Evaluation for new learning contexts how can it be fit for purpose ? *Reflecting Education*, 4(1), 127–138.

- Denzin, N. K., & Lincoln, Y. S. (2005). *The sage handbook of qualitative research* (3rd ed.). Thousand Oaks, CA: SagePublications Inc.
- EUA (2009) Improving quality, enhancing creativity: Change processes in European higher education institutions: Final report of the quality assurance for the higher education change agenda (QAHECA) project. European University Association.
- Felder, R. (1992). What do they know anyway? *Chemical Engineering Education*, 26(3), 134–135.
- Felder, R. M., & Brent, R. (2009). Active learning: An introduction. *ASQ Higher Education Brief*, 2(4), 1–5.
- Felton, J., Mitchell, J., & Stinson, M. (2004). Web-based student evaluations of professors: The relations between perceived quality, easiness and sexiness. *Assessment & Evaluation in Higher Education*, 29(1), 91–108.
- Gahan, C., & Hannibal, M. (1999). *Doing qualitative research using QSR NUD* IST* (1st ed.). London, UK: SagePublications Ltd.
- Jara, M., & Mellar, H. (2010). Quality enhancement for e-learning courses: The role of student feedback. *Computers & Education*, 54(3), 709–714.
- Leony, D., Pardo, A., de la Fuente Valentín L., de Castro, S.D., Delgado Kloos, C. (2012). Glass: a learning analytics visualization tool. In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM, pp 162–163.
- Litzinger, T., Lattuca, L. R., Hadgraft, R., & Newstetter, W. (2011). Engineering education and the development of expertise. *Journal of Engineering Education*, 100(1), 123–150.
- Marsh, H.W. (2007). Students evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In: The scholarship of teaching and learning in higher education: An evidence-based perspective, Springer, pp 319–383.
- Martínez-Monés, A., Gómez-Sánchez, E., Dimitriadis, Y. A., Jorrín-Abellán, I. M., & Rubia-Avi, B. (2005). Multiple case studies to enhance project-based learning in a computer architecture course. *IEEE Transactions on Education*, 48(3), 482–489.
- Martínez-Monés, A., Dimitriadis, Y., Gómez-Sánchez, E., Rubia-Avi, B., Jorrín-Abellán, I., & Marcos, J. A. (2006). Studying participation networks in collaboration using mixed methods. *International Journal of Computer-Supported Collaborative Learning*, 1(3), 383–408.
- Moura, I. C., & van Hattum-Janssen, N. (2011). Teaching a CS introductory course: An active approach. *Computers and Education*, 56(2), 475–483.
- Moursund, D. (1999). *Project-based learning using information technology* (1st ed.). Eugene, OR, USA: ISTE.
- Oakley, B., Felder, R. M., Brent, R., & Elhadj, I. (2004). Turning student groups into effective teams. *Journal of Student Centered Learning*, 2(1), 9–34.
- Overall, J., & Marsh, H. W. (1979). Midterm feedback from students: Its relationship to instructional improvement and students' cognitive and affective outcomes. *Journal of Educational Psychology*, 71(6), 856–865.
- Pardo, A., Estévez-Ayres, I., Basanta-Val, P., & Fuentes-Lorenzo, D. (2011). Course quality improvement using mid-semester feedback. *International Journal on Technology-Enhanced Learning*, 3(4), 366–376.
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231.
- Sheppard, S., Macatangay, K., Colby, A., & Sullivan, W. M. (2008). *Educating engineers: Designing for the future of the field*. San Francisco, CA, USA: Jossey-Bass Inc Pub.
- Steward, B. L., Mickelson, S. K., & Brumm, T. J. (2005). Continuous engineering course improvement through synergistic use of multiple assessment. *International Journal of Engineering Education*, 21(2), 277.
- Stewart, J., Oliver, W. I. I., & Stewart, G. (2013). Revitalizing an undergraduate physics program: A case study of the University of Arkansas. *American Journal of Physics*, 81(12), 943–950.
- Sursock, A., & Smidt, H. (2010). *Trends 2010: A decade of change in European higher education*. Brussels, Belgium: EUA publications, European University Association.
- Takriff, M.S., Abdullah, S.R.S., Mohammad, A.B., Anuar, N. (2011). Students' feedback in the continuous quality improvement cycle of engineering education. In: Global Engineering Education Conference (EDUCON), 2011 IEEE, IEEE, pp 374–377.
- Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times* (1st ed.). San Francisco, CA, USA: Jossey-Bass.
- Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological bulletin*, 63(6), 384.
- Wachtel, H. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education*, 23(2), 191–212.

- Walker, A., & Leary, H. (2009). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-based Learning*, 3(1), 6–28.
- Weinberg, B. A., Fleisher, B. M., & Hashimoto, M. (2007). Evaluating methods for evaluating instruction: The case of higher education. Tech. rep., National Bureau of Economic Research
- Yao, Y., & Grady, M. L. (2005). How do faculty make formative use of student evaluation feedback?: A multiple case study. *Journal of Personnel Evaluation in Education*, 18(2), 107–126.