Constantino Grau Turuelo, Oscar Banos García,
Cornelia Breitkopf

thermoE^{int}: E-Assessments for International Students in Mechanical Engineering

Using Technical Thermodynamics as Pilot for Mathematics Oriented Subjects at the TU Dresden

1 Introduction

In recent years, the development of e-learning content has proved to be a very practical tool for improving existing learning options for students and workers, as e-learning can be used anytime and anywhere. At the universities, this versatility simplifies the process of reaching a larger number of students, especially those who have difficulties understanding the lecture. This is especially true in the case of foreign students, as language is an important barrier to overcome in early stages of the study. In order to improve the integration and success of foreign students at the Technical University of Dresden, the project thermoE^{int} (thermodynamic e-assessments for international students) aims to offer online electronic exercises in English for the fundamental subject of technical thermodynamics as a prototype for other subjects of natural sciences, technology, engineering and mathematics (STEM, in German: MINT).

This contribution presents our methodology for establishing a comprehensive didactic relation of e-assessments coupled with the lecture content. The implementation of a dynamic system for a continuous year-to-year improvement through the e-learning content and the student feedback is also discussed together with the outcomes of such implementation as a starting point to the successful implementation of e-learning activities in the STEM field of study.

2 Objectives

In the winter semester of 2016, the Technical University of Dresden hosted 4827 international students (13,4 %) from 125 nations. Statistics have shown similar percentages of foreign students during the last years (TU Dresden, 2016). It also counts with numerous international programs which allow students to undertake part of their studies abroad (TU Dresden, 2017). Every year, graduate students find themselves in a more internationalized market and they demand education programs that can provide an adequate preparation. This demand especially affects the faculty of mechanical engineering, which is the largest faculty of the TU Dresden, due to the technical nature of the subjects, the global number of students in comparison to other learning areas, and the cooperation with various international universities such as the Boston University, the University of Shanghai or other European institutions included in the Erasmus program.

One specific case scenario could be the fundamental subject of »Technical Thermodynamics« where a large number of students attend and must pass the subject every year as a prerequisite to continue their studies. In the winter semester of 2016, there were 740 students enrolled in this course. Due to the number of students, it is not possible to provide much of an individualized attention and, therefore, other types of teaching methods must be developed. E-learning material can be used as a helpful tool to satisfy further learning options due to its broader range of reach and the possibility of using them at any time and any place. As foreign students can have more problems communicating their difficulties, e-learning material can ensure that the main idea of the subject can be transmitted to the students, so that there is great potential here for ensuring their academic success. The course of thermodynamics is frequently one of the subjects that suppose a greater challenge for students of engineering. Previous projects have attempted to improve the experience of teaching and learning thermodynamics and many of them have based their methodology on the use of computers and multimedia (Mulop, Yusof & Tasir, 2012).We also have previously experienced the positive effect of e-learning in the lecture of technical thermodynamics (Breitkopf, Köhler & Kretzschmar, 2014; Freudenreich, Kretzschmar & Breitkopf, 2016). This method provides a versatility that is difficult to achieve by traditional means as it comes to availability in time and place.

It has also been also demonstrated that dropouts often occur during the first semesters. According to a study by the »Stifterverband«[1] for the German Research (Donors' Association for the Promotion of Sciences and Humanities in Germany), STEM subjects lose almost 40 percent of their first-year students on their way to graduation (Berthold, Jorzik & Meyer-Guckel, 2015). This change or discontinuation of studies occurs frequently due to performance problems or inadequate study conditions. The common dropout reasons are related to stress or high pressure due to performance requirements, comprehension problems due to a large amount and bad management of the learning material, lack of practical relevance of the lecture, or even the lack of individual attention, i. e., anonymity at the university. All these problems are expected to increase for foreign students where language issues and barriers must be added, causing them to be a specially affected group. In this context, the TU Dresden recommends the development of new instruments for improving and evaluating the study quality. Increasing the students' success rate of the universities of Saxony is a declared objective of the Ministry of Science and Arts (SMWK) for 2020 (SMWK, 2011).

1 »Stifterverband is a joint initiative started by companies and foundations – the only one in Germany to be devoted entirely to consulting, networking and promoting improvements in the fields of education, science and innovation« (Cited from the homepage of the Stifterverband: https://www.stifterverband.org/english [10.12.2019]).

Since 2015, the TU Dresden has done a great effort to introduce e-learning methods in the university. This decision intends to orientate teaching strategies towards a digitalization that opens the university to more international learning and collaboration strategies. Hence, e-learning makes use of the advantages of information and communication technologies to support the teaching and learning processes. Of special importance are the possibilities of creating a more interactive and individualized framework for the students. These features are especially valuable in the course of technical thermodynamics due to its versatility and the different background of the students.

Within this frame, the aim of thermoE^{int} is to develop, adapt, and establish subject-specific e-assessments with mathematical impact in addition to the regular teaching material. This is specially intended for international students of mechanical engineering at the TU Dresden. So far, there are no comparable offers with a subject-specific background, especially for the fundamental subject of thermodynamics, which, in addition to mathematics and mechanics, is decisive for the success of studies in mechanical engineering. The e-assessments are not intended to be language tests but subject-specific tests so as to learn and use technical terms required to work in any thermodynamic field. This should reduce the time required to understand the subject due to lower time consumed in activities not related to the subject itself such as the direct translation of the learning material. It has to be emphasized that German courses enable a basic knowledge in general language at the beginning but this previous knowledge of general language is not sufficient to follow a specialized lecture (a specialized language German course does not exist). An additional English-language e-assessment with a subject-specific orientation could close this gap (Rech, 2013).

The proposed self-assessment material should provide resourceful information for the international students to prepare themselves for the lecture, seminars, and the final exam while decreasing the language barrier that they must overcome. This is also useful for German students as a preparatory course for a later stay abroad in any of the existing exchange programs of the TU Dresden with other international universities or to be better prepared for an international career. However, the main target group is the foreign students as they have a higher trend, in comparison to German students, to drop out of higher education: the graduate rate of foreign students in German studies is only of the 40 to 50 percent compared to the 70 percent for German students. In order to improve the academic success of these students, it is necessary to create adequate support services at the universities (Rech, 2013).

In summary, the project pursues the following sub goals:

- the development of a thermodynamic, mathematically-oriented catalogue of online tests in English in the form of e-exercises which allows the students to practice and check more complex learning scenarios as well as to develop their cognitive processes, while decreasing the language barrier. In the lecture Technical Thermodynamics, the students are given content orientation, so that they can solve the available competence-oriented, didactically conceived tests.
- a feedback function is integrated to receive detailed feedback on the results from the students. This leads to the development of better commented solutions to clarify the possible gaps in knowledge that the students may have and discuss them in the corresponding seminars. The concept is borrowed from the so-called blended learning concept.
- development of online examination scenarios for an intensive preparation of the final exam so that the results can be improved.

The implementation was carried out in OPAL (ONYX) (BPS Systems, 2019a; BPS Systems, 2019b) due to its suitability for the proposed project and its implementation in Saxony as the standard education platform.

This work presents the methodology developed by the Chair of Technical Thermodynamics, part of the Faculty of Mechanical Science and Engineering at the TU Dresden, for establishing a comprehensive didactic relation of e-assessments coupled with the lecture content. The implementation of a dynamic system for a continuous year-to-year improvement through the e-learning content and the student feedback is also discussed as well as the outcomes of such implementation in the scope of the lecture of »Technical Thermodynamics« and »Energy and Thermodynamics«.

3 Building e-assessments: Methodology

Before building the e-assessment content, it is necessary to establish the appropriate requirements to fulfill the objectives. The methodology employed is based on experience. Teaching is an area that cannot only be evaluated with quantitative data due to its dependence on human understanding capabilities and the cognitive processes. As a consequence, qualitative data such as oral feedback or interviews is fundamental to understand the necessities of the students and the problems they face at a daily basis. The most valuable feedback and ideas are provided by the teaching staff as they are continuously in direct contact with the students. They are able to track the advances, reactions, and the issues that the students may have under different educational situations.

Before constructing suitable e-learning exercises and setting their objectives, it is crucial to gather experiences and opinions to establish the most important questions which should be addressed and convert them into the necessary didactic and technical specifications. A flow diagram exposing this conversion process is shown in Figure 1.

The key factor to success in the implementation of any teaching strategy is the education experience. As a result, the most important source of data collection comes from the educational staff. Qualitative data are collected through interviews and conversations with the responsible Professor of the subject, the seminar leaders, tutors

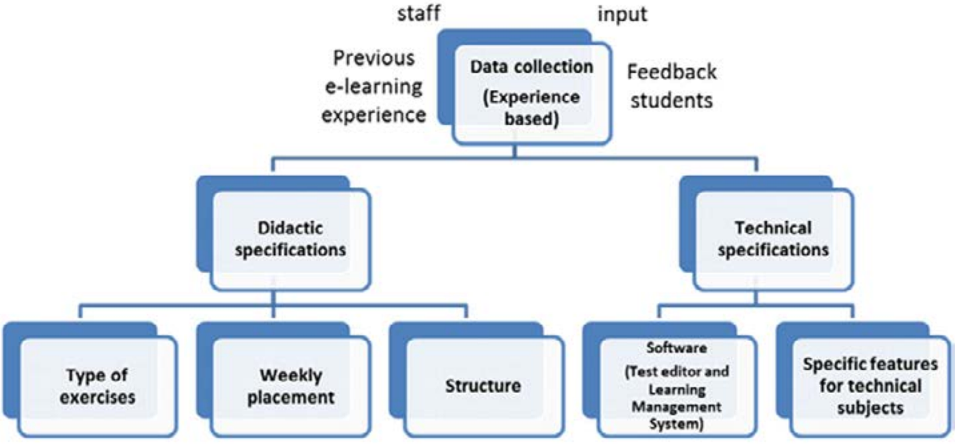

Figure 1: Workflow of the e-assessment planning process

and other members of the chair. From the Professor, it is possible to determine the structure of the subject and the most frequent inquiries asked by the students after each lecture. This knowledge is then used to develop conceptual exercises based on the theoretical background of the subject. Tutors and seminar leaders can collect data about the students' current solving competences which can be gathered from seminar leaders and tutors. The information is then utilized to develop more practical online exercises.

The feedback from the students is also of central importance. This feedback can be collected directly by electronic or personal means. The personal form would be through a direct informal conversation. The second option would be to receive their feedback electronically if an e-learning system has already been implemented. Fortunately, this project was cemented on previous project experiences, so that a lot of the required feedback to start the e-learning development was already available. Those projects and their respective developments were: thermoE for evaluating the theoretical, conceptual and didactic framework for the creation of thermodynamic e-assessments (Breitkopf et al., 2014; Freudenreich et al., 2016), thermoSA for the practical application of thermodynamic self-assessments (Breitkopf, 2015) and SPAts for the development of self-/peer-assessments in supply chain management and thermodynamics (Freudenreich & Lorenz, 2015). After successful implementation of these projects, the already existing e-assessments have been already tested and evaluated by about 500–600 students each year. Those electronic assessments were created within the scheme of the subject »Technical thermodynamics«. This gave us about 10–50 comments from students per online test, which aided to the implementation and improvement of new and existing exercises.

The final source of gathered information was other external resources, such as conferences, literature, questionnaires, surveys, and experience from other departments. In this context, 22 students from abroad participated in a questionnaire as a prerequisite for the development of thermoE^{int}. The objective was to know what their expectations from an e-learning platform were, what their skills were and which suggestions they had. Among the results, it was found out that 85 percent of the students who participated in the questionnaire spent the same or more time in self-study in comparison to the time employed in lectures/seminars, 62 percent used Internet resources frequently to study, and between 60–70 percent of the participants used Internet to download technical materials. These quantitative data show that most students are already familiar with online resources. As a result, approaching them through the use of e-assessments as the learning tool is a feasible method.

After collecting the data, problems, requirements, and solutions will be discussed. These requirements are divided into two categories: didactic and technical specifications.

3.1 Didactic specifications

Overcoming the forgetting curve (Figure 2) is one of the main objectives of a learning strategy. The learning curve shows the retention of information during a lecture. A few hours later, in the absence of a review of the lecture content, there is a decrease of the retention along the first forgetting curve (Averell & Heathcote, 2011; Ebbinghaus, Ruger & Bussenius, 1913; Murre & Dros, 2015). The utilization of e-assessments and seminars help to minimize the retention problematic. The previously acquired knowledge is refreshed and follows the other corresponding forgetting curves. The chances of forgetting the content of the lecture is minimized as the number of teaching events increases.

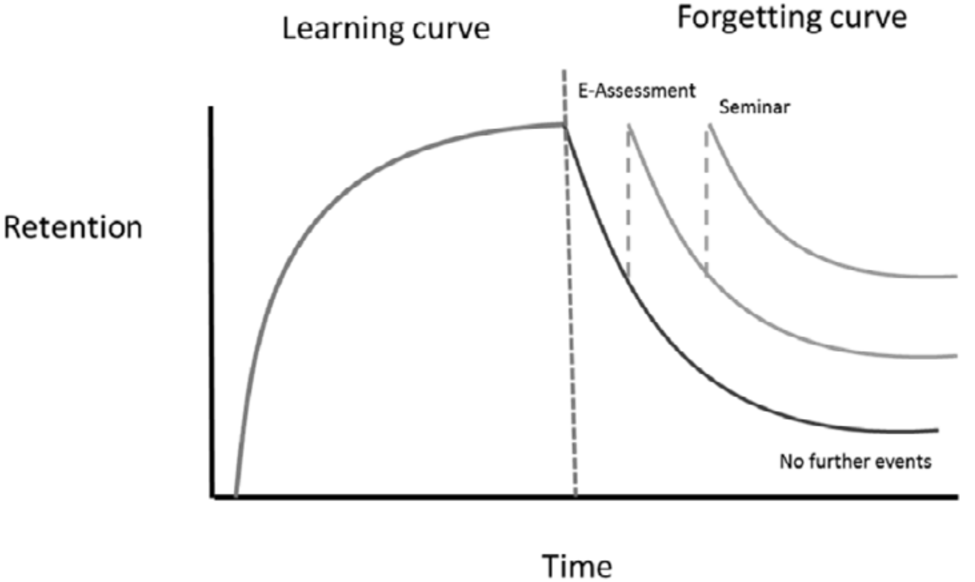

Figure 2: Representation and adaptation of the learning and forgetting curve

In order to overcome the forgetting curve, it is important to integrate the tests within the structure of the lecture. In this case, the taken approach was to divide and distribute the tests per lecture week so that students can review and do e-assessments related to the content of the last lecture. There were twelve lecture weeks in the subject »Technical Thermodynamics«, therefore, twelve e-assessments were prepared. On the day of the corresponding lecture, the respective test was unlocked. The students had access to the test throughout the week until the next lecture. The number of attempts was restricted to two in order to force an active learning effort so that »trial and error« techniques are avoided.

In each test, the complexity of the exercises was progressively implemented. The simplest tasks were placed first with key concepts for its later use in more complex tasks. The allocation of different task types is also crucial. The tasks were divided into two forms: conceptual tasks and solving skills tasks. The conceptual tasks are intended to encourage the students to learn the concepts and theory that are conveyed in class. The second objective is to familiarize the student with the didactic material, i. e., to make them to look for the expression they must use and learn, and where to find it. This was accomplished by defining gap text exercises where the student had to make them look for the number of the requested equation in the manuscript/formulary or the page where it is located and write it in the gap. This was done to improve the lecture material management skills.

Solving skills tasks are mainly mathematical exercises in which students must solve a numerical problem. For technical studies, most exercises belong to this classification. Solving skills exercises promote problem-solving competence, logic, and argumentation. It is advisable to mix conceptual tasks with calculation tasks, since they contribute to the motivation of the student due to the variability of the exercises. This tries to avoid boredom in the long term. In the first e-assessments of the semester, the percentage of conceptual exercises should dominate as they are crucial to understand the basics of the subject and increase the handling ability of equations, tables, and other similar material. The balance should be gradually shifted to solution-oriented exercises as the course approaches the last weeks of the subject. In fact, the planned structure in the project thermoE^{int}, the first test had an 86 percent of short conceptual tasks (from a total of 29 tasks), while, in the last test, this number was reduced to 12.5 percent (from a total of 24 tasks). The shifting of exercise load was done to facilitate the practical preparation for the exam. Furthermore, the e-assessments should follow a logical step-by-step solution structure so that students comprehend how to solve typical exercises of the subject.

The last and most important required didactic specification of our project was to help foreign students become familiar with the subject. Let us first present some quantitative data from the international student survey mentioned before: 52 percent of the students who participated in the survey had language difficulties during their studies, while 71 percent confirmed that they had to overcome a major barrier to study due to language. The first idea to support international students was to offer exercises in English only. The real obstacle, however, was to transfer the technical knowledge acquired in German into English e-assessments and vice versa. The found solution was to translate the most important tasks so that the students can see the connection between the German and English technical terms and prepare them for an interchangeable use.

A tutorial was also developed to improve their adaptation to the online e-evaluations from the very beginning. In this way, they can quickly understand how the e-learning system works. According to the feedback of the students, the tutorial was a much-appreciated resource.

3.2 Technical specifications

The technical requirements depend to a large extent on the software, learning management system or platform, and associated functions. In thermoE^{int}, the main software prerequisite is to have as many exercise customization possibilities as possible. Regarding the learning management system (lms), it must be easy to use, user-friendly, and transparent. ONYX Editor (BPS Systems, 2016b) was the program chosen for the development of the e-assessments as it offers a wide range of options and is very well integrated in OPAL, the standard lms established in Saxon universities. The reason of this choice is that the students of our university are already acquainted to OPAL, so it is natural that they already use it in most degree subjects. In addition, ONYX offers many tools to enhance the e-learning experience, allowing us to have more flexibility when making the content of the e-assessment.

The first technical requirement comes from the didactic specifications: the students should learn the logical structure to solve every practical exercise type. Therefore, it is important to provide a step-by-step solution so that students know what the first step to get to the right solution is and what the solution procedure is. To do this, the ONYX capability of creating exercises with sections was employed. Section-based exercises present different subtasks. Each subtask is presented sequentially but not simultaneously, i. e., subtask b is only shown after subtask a has been completed. The main text is shared with all the subtasks. In the main text, the student can find the formulation of the overall exercise and all initial data, so that the student does not need to go back and forward to review the initial information. The main text is always accessible and visible while being in any corresponding subtask (see Figure 3). The second reason for using this structure is to avoid subsequent errors. Each subtask is evaluated before moving on to the next one. If the solution is incorrect, the interface shows the right solution and the students can thus utilize this value in the following subtasks (see Figure 3). With more common approaches, an error in the beginning of the exercise would lead to a false solution of the rest of the subtasks after a big amount of calculations. The method presented here aims to preserve the motivation of the students due to the consideration of subsequent errors and enforces the students to learn the logical solving procedure steps.

Real-time feedback is also a recommended feature that should be present throughout the section structure, as shown in Figure 3. The ONYX Editor allows you to define a text that provides sub-task-based feedback for students. If students have solved a subtask incorrectly, a text appears with instructions on how to solve the subtask correctly, for example, specifying the subject’s script page, the formula page, a reference book, or, if the subtask is particularly difficult, the entire

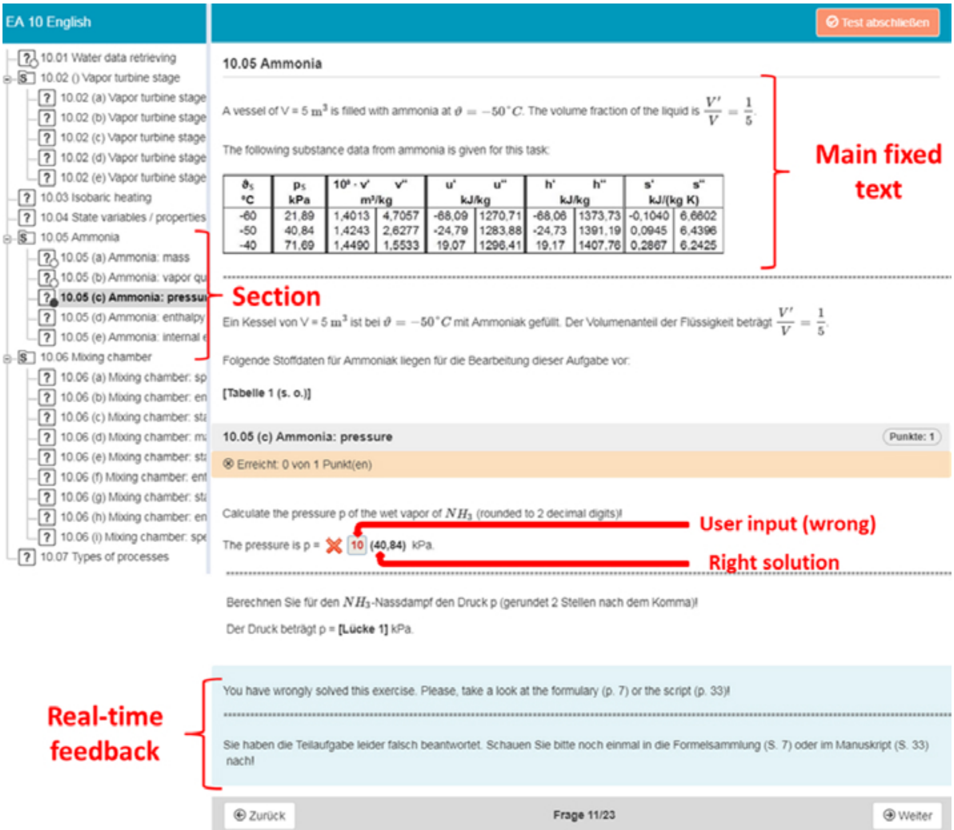

Figure 3: Screenshot of one test after a wrong input

resolution process. With this methodology, students can actively seek the solution, increase the learning experience and immediately learn what they have done wrong. They also learn how to use didactic material more efficiently.

In STEM subjects, students will have to do many calculus exercises to improve their problem-solving skills. This can also be practiced by using e-assessments. However, due to the use of different decimal numbers or slightly different parameters, each student frequently obtains a different solution for a particular task. In order to avoid this, the definition of some number tolerances must be implemented. In the project thermoE^{int}, a common used relative tolerance is defined between 1 and 5 percent after having analyzed the effects of a solution on the extreme sides of the tolerance. The chosen tolerance must also prevent erroneous solution methods from being mistakenly regarded as correct. For text gaps exercises, a tolerance has also been introduced so that a misspelled letter is not considered incorrect.

As mentioned in the didactic section, a large number of exercise types avoid boredom and increase the motivation of the student. It is recommended to combine as many exercise forms as possible. In this project, the following types of exercises were implemented (BPS Systems, 2016a; Breitkopf, Grau Turuelo & Banos García, 2017; Freudenreich, Grau, Breitkopf & Kretzschmar, 2017):

- Choice interaction: this exercise includes a list of possible answers where one or several are right.
- Text entry interaction: the task consists on completing a text where several words are missing.
- Extended text interaction: this task offers an area where the student can write free.
- Match interaction: you need to drag several items and drop them next to the right description.
- Matrix interaction: this exercise presents a table where the user has to mark for each of the elements in the first column the matching elements on the first line.
- Order interaction: the user has to order some elements by dragging and dropping.
- Hot spot interaction: the task consists on selecting the right area over an image.
- Hot text interaction: the task consists on marking the right word/s inside a text.
- Inline choice: the user has to fill gaps inside a text by choosing among several options given for each gap.
- Upload interaction: this exercise allows to upload a file.

- Numerical input interaction: the exercise contains gaps where a number is introduced and compared with a reference with a certain tolerance.
- Calculation interaction: in this case the value introduced is compared with a calculation made by the software itself according to a defined random variable.
- Formula comparison: a formula introduced in Maxima language is compared with a reference.

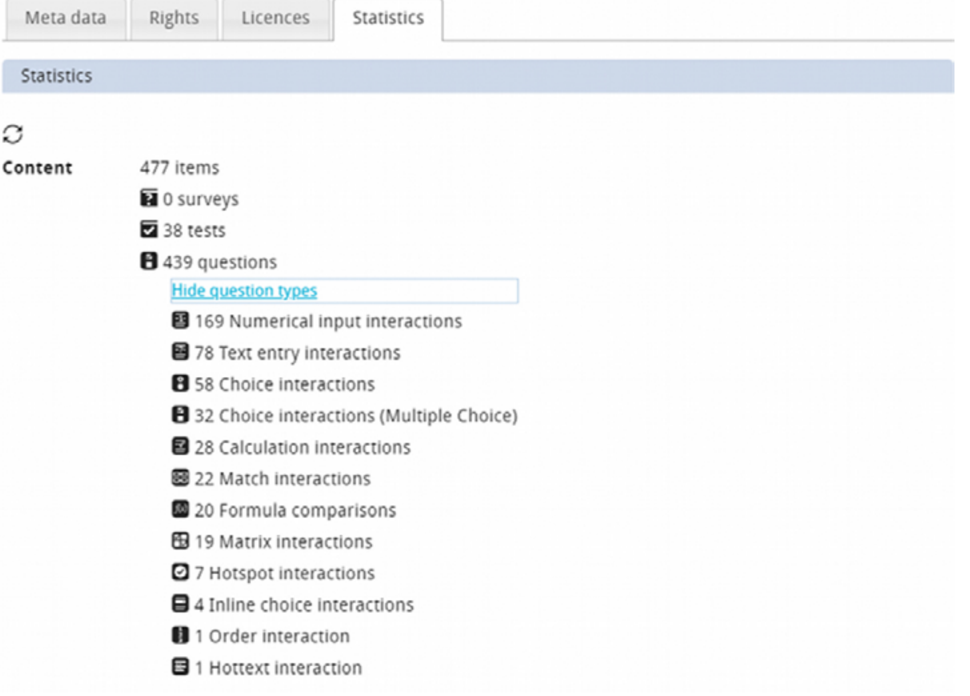

Figure 4: Registered subtasks during the project thermoEint (not all included in tests)

The duration of each test was prepared to take between one and two hours, for a total of 12 tests (12 teaching weeks) with 116 online exercises in total for the final offered format. Figure 4 represents the total developed and reworked tests and subtasks to develop such an extensive offer. Another introduced feature is that they have the option of stopping the test at any time and continuing the exercises at a later time or on another day. This gives the students a lot of flexibility, making them decide when and where they want to take the weekly scheduled electronic assessments. It is advisable that the students do exercises progressively during the lecture week, as the revision of exercises about the week topic at different weekdays reinforces the retention of the lecture content as shown in the forgetting curve (see Figure 2).

Finally, after each electronic assessment, a feedback survey was available to learn about test behavior, potential problems, and student satisfaction.

3.3 Year-to-year improvement cycle

After the creation of the first version of the e-assessments, and when it has been completely tested, new revisions should be performed in a yearly basis to refine the content of the e-assessments (see Figure 5). Depending on the timeline in which the revisions are carried out, two types of modifications can be distinguished: changes between docent years and changes during the semester.

For a dynamic refinement during the semester, a feedback survey is carried out after each e-assessment. The learners can anonymously express their opinion on the difficulty, suitability, and duration of the test. They are also allowed to give feedback on the problems they had on a free text box as well as contribute with possible suggestions for improvements or other kind of comments. This information can be even used to make changes »on the fly« if it does not affect the outcome of the weekly task or exercise. It is recommendable to read the first feedback shortly after unlocking the tests to check that everything is working correctly and, if any issue is found, prepare the changes for later e-assessments which may be affected.

Concerning the yearly improvement cycle, there are two sources of information to improve the e-assessment content: analysis of student feedback after each e-assessment and statistical data analysis. Our

Figure 5: Workflow diagram for the yearly improvement cycle

learning management system, based on OPAL and ONYX, can collect anonymous user statistics such as the percentage of tasks with correct/incorrect answers, the summary of the most frequently written values in numerical exercises and the overall pass rate of each e-assessment test. With these data, the deviations in the average results can be analyzed and the more problematic exercises can be determined. This allows instructors to make appropriate changes to the seminars/e-assessments in order to achieve a better and more consistent set of tests and thus increasing the teaching quality. Along the refinement cycle, it is also possible to implement new ideas and/or functions not available before.

4 Results

In this section, the features described in the didactic and technical specification of the creation of e-assessment content will be shown as output of the project (as the tests are, at the time of writing, still running), including some statistical data, comments of the students and some evaluation figures.

4.1 Technical outcome

As described in the section 3, the self-assessment tests were realized in the learning management system OPAL. This platform is a product of the Bildungsportal Sachsen GmbH and enables the creation of online courses. Authors have a set of tools to design different kinds of content, among them; online tests (see Figure 6). This platform was chosen because it is systematically used by universities and cooperative education institutions of Saxony, which improves the adaptation of the students. Furthermore, previous works have shown the suitability of this platform for the project objectives (Breitkopf et al., 2014).

Figure 6: View of the OPAL online course

The tool used for the edition of the exercises was ONYX. This is another web-application of the Bildungsportal Sachsen GMBH embedded in OPAL. Students have direct access to the exercises and the educational staff has direct access to the results without leaving the OPAL domain.

As explained in the specifications part of this work, the tasks can be grouped in sections (see Figure 7) where each answer can be submitted independently of the others. In this way, the user always has the main statement of the exercise visible while working on the different tasks of such exercise. After submitting numerical answers, the student receives the correct answer. The student is allowed to jump back and forth along the tasks, by using the navigation tree, so that previous solutions can be reviewed.

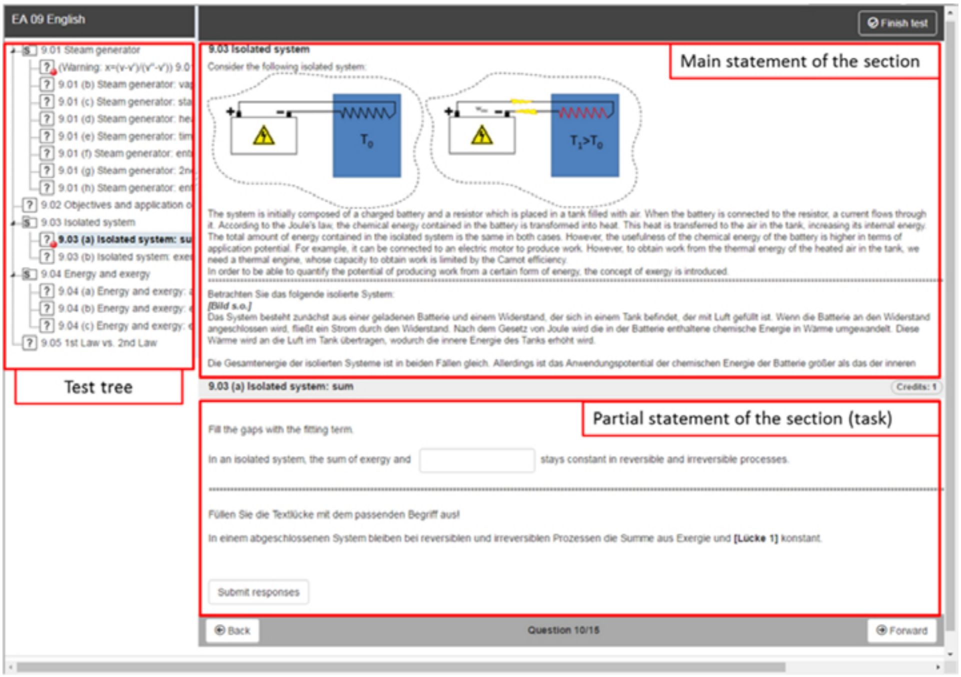

Figure 7: View of an online test

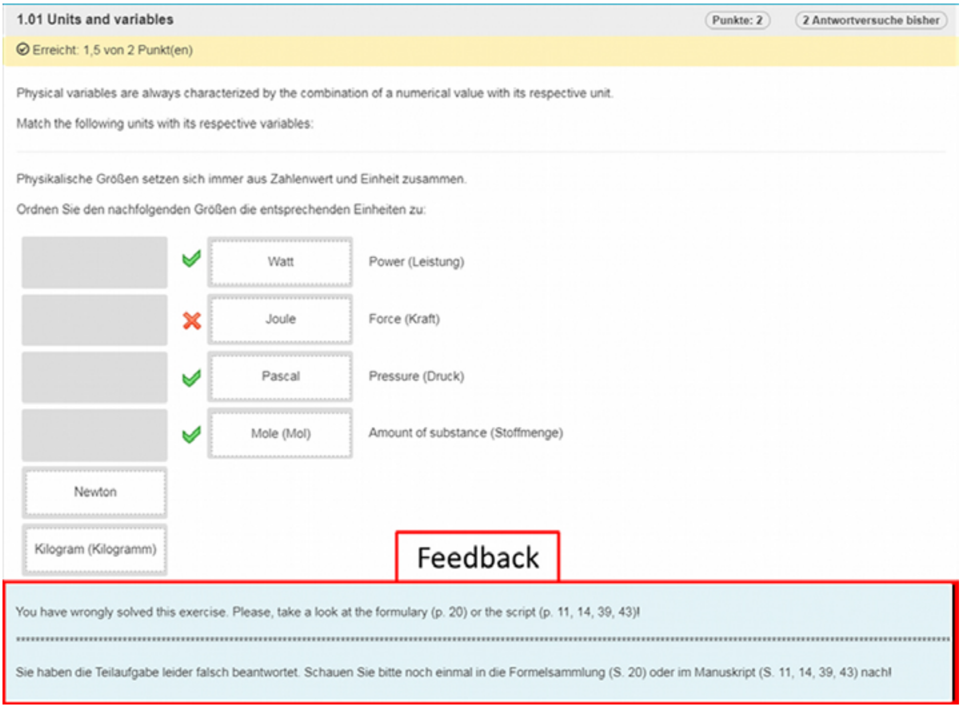

Figure 8: Feedback for the students

After submitting the answer, the student also receives a feedback message according to the submitted answer (Figure 8). This text is programed to display an encouraging message in case the student submitted the right answer and a message with a literature/ manuscript/formulary reference in case the answer was wrong. Some kinds of exercises allow even a customized feedback so that the user receives different messages depending on the made mistake. Apart from that, numerical and text tolerances, as described in the technical specifications, were correctly implemented and the time span required to complete each test was not limited (apart from the weekly deadline). The students could successfully stop the test and continue later without losing the previous work. All technical specifications are working properly and only small incidents were reported.

A feedback survey is available together with each test. In this survey the student can rate the test in terms of difficulty, time consumption, suitability, common problems and so on. It also provides three free text areas where the student can write suggestions, complaints and comments. These surveys are anonymous and completely voluntary.

4.2 Pedagogical outcome and impressions

The objective of lowering the language barrier in STEM subject was attempted by the creation of a comprehensive relation of conceptual exercises in both German and English languages. Such exercises were designed in ONYX and were later used to create online tests in the learning management system OPAL. The exercises are a set of theoretical and practical problems whose statements are presented in English and German. In this way, the students can see the correlation between the concepts and get acquainted to the technical terms on both languages.

With the implementation of this relation of exercises, we have overcome certain challenges that have contributed to the future introduction of e-learning activities in subjects in STEM fields, an area where the utilization of this kind of methodologies is not as extended as in social sciences (Breitkopf et al., 2014). The new division of the exercise in tasks and the correction of the submitted answer prevented the consecution of errors which is a significant advance to implement the suggested e-assessment approach in other technical or mathematical subjects. Another proposed approach to learn how to interpret diagrams can be observed in Figure 9, something which is not commonly easy to evaluate in online exercises.

In the following, some reactions of the students from the feedback system implemented in the tests will be shown in relation to the different taken approaches.

From the data gathering and the planning steps of the project, we observed that the learning curve of the learning management

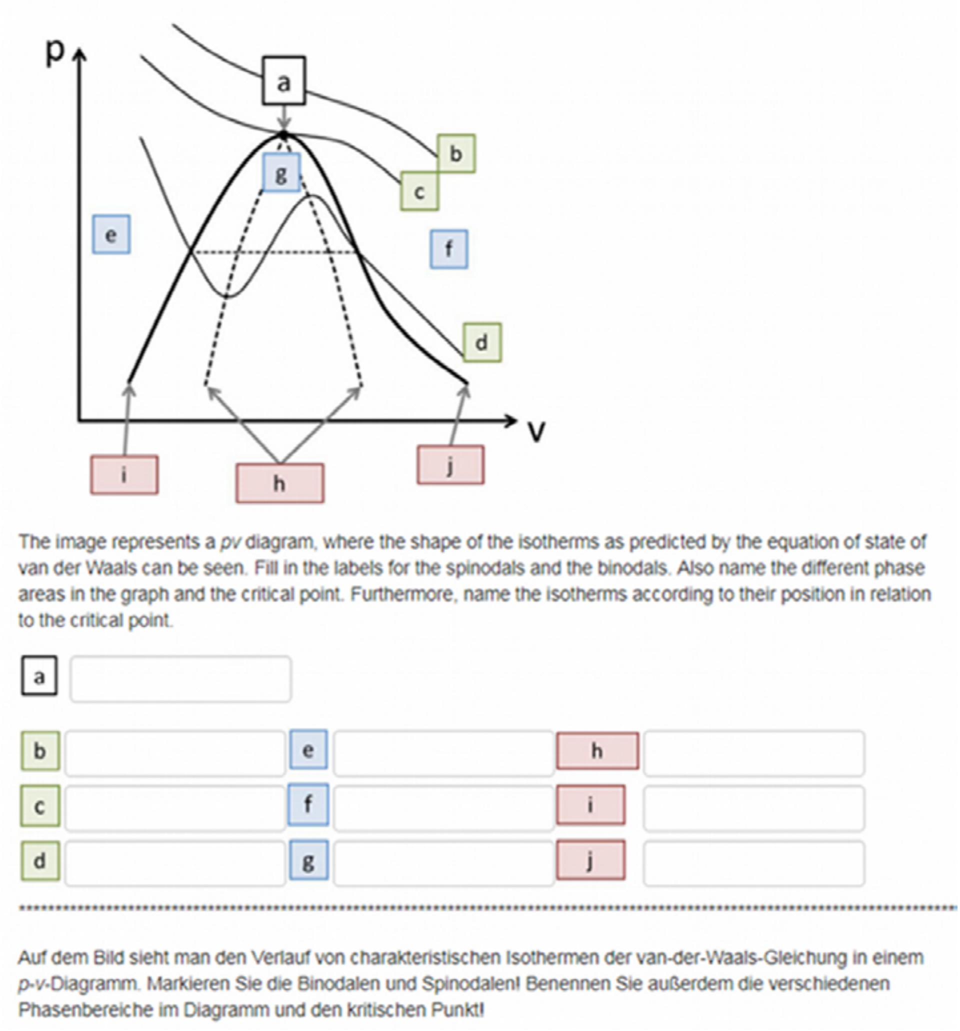

Figure 9: Diagram interpretation as example for technical thermodynamics

system was sometimes the reason why some students dropped the online tests. To lower the learning curve, a tutorial-test was created with simple exercises, which could be answered without previous knowledge of thermodynamics, but helped the user to familiarize with the different kinds of exercises and the features of each of them.

This was really welcomed by the students: »I like the lovely designed examples«.

Another of our objectives was to encourage the student to get used to the learning material: formulary, lecture script and books. This was achieved directly by asking the students to introduce the number of the equation of the subject formulary that must be applied in a certain exercise. An indirect way of making them acquainted of the learning material, which turned to be really effective, was the feedback messages of the exercises: »The tips to the related pages in the formulary were really helpful!«; »Indeed, the hints were usually helpful«; »The references to pages of the script by wrong answers are super!«

By splitting the exercises in different tasks and steps, we obtained two different outcomes. On one side, we can guide the reasoning of the student so that, from a particular example, the general resolution methodology can be understood, increasing the solving skills of the user. On the other side, solving-skills exercises can be complemented with conceptual exercises. This complementation of the application-based exercises with the theory-based exercises can be done in two structures. The first option is to start with a set of conceptual exercises that finishes with a mathematical example. This is the structure that we typically used in the first lessons of the course since their content was more theoretical. The second option is to alternate mathematical exercises, where we intercalate theoretical with applied questions, with the concepts required to solve the next mathematical step. This structure is the one that we typically use in advanced lessons because they are more application-based. The right combination of both kinds of exercises can help to see the relation between theory and practice as well as preventing the tediousness of the tests.

The exercises were, in many cases, accompanied by explanatory pictures (see Figure 10). These pictures were intended to make a clearer explanation of the exercise. In the literature, it is possible to find results that prove that the appearance of the exercises has an influence on the engagement of the student (Mulop et al., 2012).

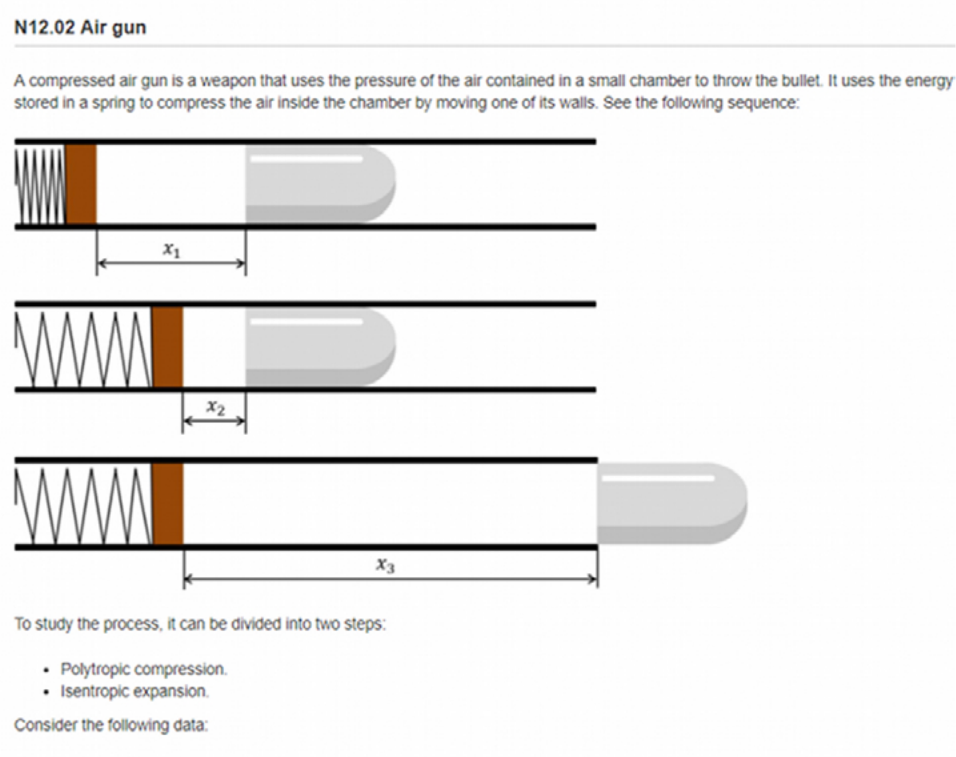

Figure 10: Explanatory picture from an e-assessment example

This was indeed the outcome observed in our project: »funny comics«, »Overall, I liked the test very much. I am very impressed by the fact that so much effort is put into providing us with an optimal as well as easy and fast preparation of the material. A big thank you from here!«

The exercises were released on a weekly basis. The content of the tests was related with the theoretical content treated in the lecture during the previous week. Therefore, the objective of the exercises was to help to the preparation of the exam, not only by training with similar exercises, but also by allowing the students to self-assess their own knowledge. The tests also played the role of partial exams to evaluate the students' level.

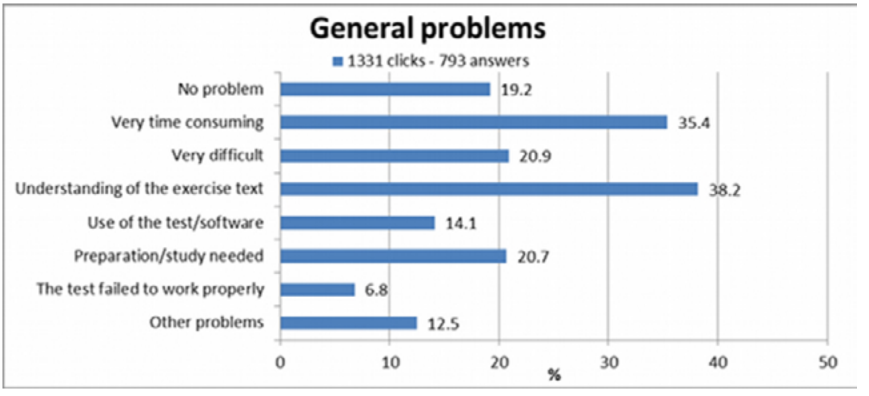

Figure 11: Survey stats as clicked by the students, average of all e-assessments

The voluntary surveys at the end of each e-assessment, as explained in the section 4.1, turned out to be a very useful tool to improve the communication with the students. They provided statistical data that allowed a quick determination of failures in the design of the tests (see Figure 11). For instance, for the following year after these data was collected, the amount of exercises was reduced to be one hour or less per e-assessment. A review of the most problematic exercises was also done to modify the exercise test to make it easier to understand.

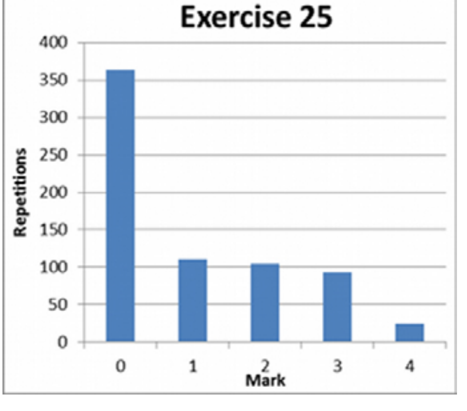

Figure 12: Exercise stats. X-axis shows the number of obtained points until the maximum defined by the exercise (4 points maximum in this example). Y-axis is the number of trials which scored the corresponding mark

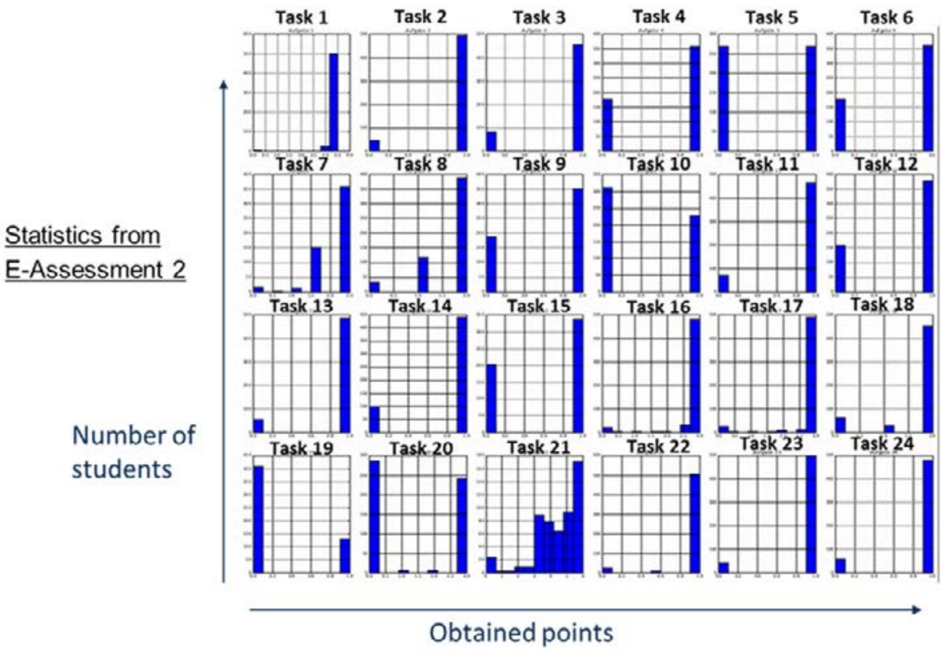

Figure 13: E-assessment stats, exercise per exercise X-axis shows the number of obtained points until the maximum defined by the exercise. Y-axis is the number of trials (students) which scored the corresponding mark

The free text introduction was used for many purposes. Firstly, it was used to upload questions that were answered in the following seminar, for example: »Please repeat the calculation of entropy / enthalpy in the tutorial.«, »I've had some trouble with that: 10.06 e) and at 10.06 h)«. In this way, the tutorials could be designed to better fulfill the interests of the students. Secondly, the students reported their opinion about specific details of the tests: mistakes, sentences hard to understand, typing errors, etc. Some examples of corrections noticed by the students which were rapidly solved with the students' feedback: »In task 2.12, the solution yielded a strange value. The air density under the given conditions was given as approx. 12.5 kg/m³. Comparative values, however, range between 1.2 kg/m³ and -1.25 kg/m³.«, »The following worked in the tutorial: A space between the letter and the mark: (a * b). In the test, only a*b was declared correct, not a * b as well.«. Hence, the surveys provided also valuable information for the design of future e-learning material. Of course, not all received feedback was positively oriented but, usually, the negative feedback failed to communicate what the real problem was, or it was giving the same information of the Figure 11. For instance: »This time very time-consuming. Learning effect questionable.«, »Very time-consuming. Partially some understanding problems«, »impossible«. Consequently, the data provided in the Figure 11, but for every corresponding e-assessment, revealed more productive information for the development of the tests.

TEST 2

Nr.	Aufgabe	Statistik			
1	2.1 State behavior	6 Antworten	83%	17%	0%
2	2.2 Specific gas constant	6 Antworten	83%	0%	17%
3	2.3 Molar volume of ideal gases	6 Antworten	67%	0%	33%
4	2.4 (a) Density and specific volun	5 Antworten	80%	0%	20%
5	2.4 (b) Density and specific volun	5 Antworten	60%	0%	40%
6	2.5 (a) Balloon of helium: volume	5 Antworten	100%	0%	0%
7	2.5 (b) Balloon of helium: specifi	4 Antworten	0%	0%	100%
8	2.5 (c) Balloon of helium: buoyar	6 Antworten	0%	17%	83%
9	2.6 Vapor quality	6 Antworten	50%	33%	17%
10	2.7 (a) Coffee: phase test	6 Antworten	50%	0%	50%
11	2.7 (b) Coffee: wet steam	6 Antworten	33%	33%	33%
12	2.7 (c) Coffee: temperature	6 Antworten	0%	0%	100%
13	2.7 (d) Coffee: calculation of vap	4 Antworten	25%	50%	25%
14	2.7 (e) Coffee: triple point	5 Antworten	60%	40%	0%
15	2.8 (a) Cylinder with water: volur	5 Antworten	60%	0%	40%
16	2.8 (b) Cylinder with water: phas	6 Antworten	33%	0%	67%
17	2.8 (c) Cylinder with water: press	4 Antworten	75%	0%	25%
18	2.8 (d) Cylinder with water: vapo	3 Antworten	33%	0%	67%

2.7 (a) **Lücke 1: subcooled liquid (50% | 50%)**

50%	subcooled liquid
16,7%	wet steam
16,7%	k
16,7%	saturated liquid

Figure 14: Left: statistical analysis of right answered subtasks (left), partially correctly answered subtasks (center), and wrongly answered subtasks (right). Right: statistical analysis of the given answers in a text gap

The statistical analysis of the surveys was complemented with the statistical analysis of the results from the tests. The time tracking of the learning platform allowed us to determine whether the workload of the tests was adequate according to the course program or not. Furthermore, an overview of the general obtained results in a specific question allowed to spot exercises whose complexity was over the average level of the participants or exercises whose statement could be improved (see Figure 12, 13, and 14). Studying the results, it is possible to easily spot problems. For instance, in Figure 13, the subtasks 19 and 20 (bottom left of the figure) have, in average, poor results. This is indicated by the distribution of the marks. When most of the students have a low score, the score distribution tends to go to the left side of the graph, while high scores move the score distribution to the right side. This can indicate a problem with the concept which must be employed in the task or another problem associated to the test creation and it must be revised during the year-to-year improvement cycle. In the case of text gaps, common misconceptions can also be observed (Figure 14 – right)

In personal interviews with the participants of the ThemoEint project, we could see that the offer of bilingual exercises and the new carried out strategic measures helped to lower the language barrier. The effects were especially noticeable during the final exam when the students admitted that they found no issues in the understanding of the exercise formulation.

Table 1: Statistics obtained from the course of »Energy and Thermodynamics«

	Percentage of correctly answered subtasks	**Percentage of partially correct answered subtasks**	**Percentage of wrongly answered subtasks**
Summer Semester 2016 ($N_{project}$=15)	50.5 %	17.9 %	31.6 %
Summer Semester 2017 ($N_{project}$=7)	74.0 %	11.0 %	15.0 %

Statistics were collected from two courses where the new e-assessments were implemented. The first group was formed only by foreign students of the subject »Energy and Thermodynamics«. The results of the tests are enclosed in Table 1. These data were collected by performing the same analysis as in the Figure 14 – right for every test and computing the average for all tests in the subject. The group of 2016 tried a first version of the e-assessments, while the group of 2017 tested the first improvement cycle. All students registered in the course did the e-assessments. As it can be observed, with the new e-assessments, the percentage of right answers was increased by a 46.5 percent, while the wrong answers were reduced by a 52.5 percent. That is a noticeable improvement from the new e-assessment structure. The students who took part in the subject were all new in 2016 and 2017.

Table 2: Statistics obtained from the course of »Technical Thermodynamics«

	Percentage of students who did e-assessments	Percentage of students who passed the e-assessments
Winter Semester 2015 ($N_{not\_in\_project}$=600) (N_{total}=825)	72.7 %	97.2 %
Winter Semester 2016 ($N_{not\_in\_project}$=500) (Nproject=12) (N_{total}=740)	69.2 %	98.0 %

In the course of »Technical Thermodynamics«, due to the amount of students (600 in 2015 and 512 in 2016) who participate in these tests, statistical data cannot be so much detailed. In 2015, the project had not started yet so not so meaningful statistical data could be recovered. In 2016, more detailed data could be collected but technically, there were limitations, as some of the data had to be collected at per student basis. For small groups, this is possible to do, while for so large groups, that is not feasible. For this reason, it is possible to compare how many passed, in average, the tests, as detailed in Table 2. Furthermore, the data from 2016 is separated in two: those who did the English test and were included in the project (12 students) and those who only benefited from structural changes (500 students). As there is no English group comparison possible from 2015, the effect seen in Table 2 is a mix of the effect of the language measures and the new test structure, while the latter should be predominant due to the difference of number of students. The participation rate is lower due to the fact that some students of 2016 did already e-assessments in 2015 and they did not repeat the experience. The passing rate is slightly higher but it is difficult to draw a clear conclusion as the passing rate was already quite high in 2015. Unfortunately, the same statistics as the smaller »Energy and Thermodynamics« group cannot be obtained due to technical limitations.

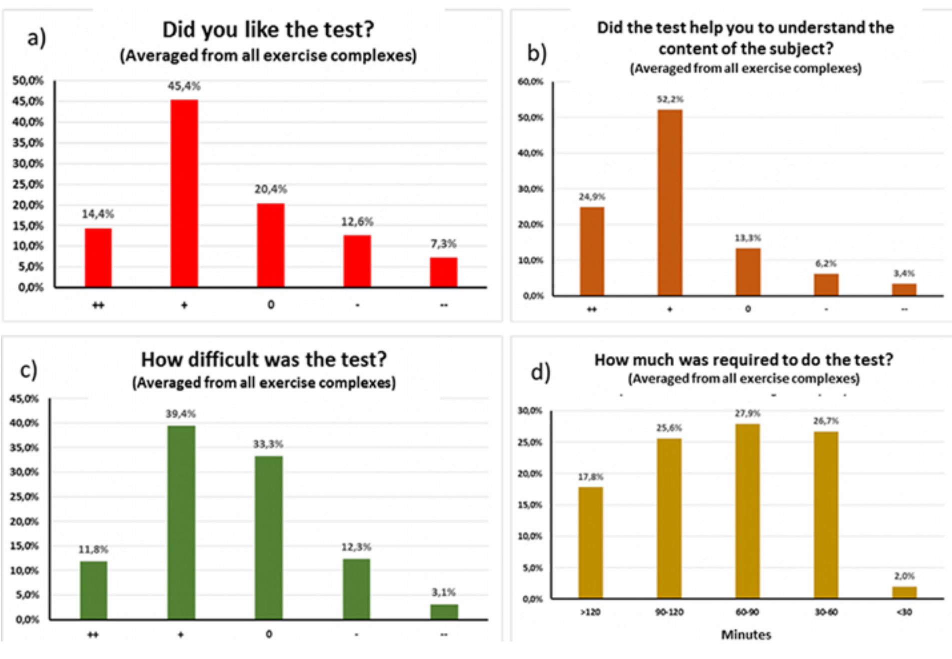

Figure 15: Feedback survey global statistics (project and no-project members) from 796 answers in total, 66.33 answers in average per e-assessment

As the last analysis, the results of the feedback survey of the tests were collected to control if everything was working as expected. It can be observed in Figure 15a that the majority of students were satisfied with the e-assessments. Figure 15b shows that more than the 75 percent of students affirm that it was quite useful or very useful to understand the content of the subject. The difficulty was perceived as regular or a bit high. This, in combination with the statistical data of the Table 1 and 2, is a good sign as even when they did not perceive the tests as simple ones, the pass rate has generally increased. In other words, the new e-assessment structure is working as intended. Finally, Figure 15d shows the time range between the defined limits (1–2 hours). However, there is a significant amount of students who affirm that it took more than what initially planned so a revision of the exercise length must be performed.

5 Summary and conclusions

In this work, the methodology developed during the project thermoE^{int} was shown.[2] A comprehensive relation of e-assessments was constructed for helping the foreign students to overcome language barriers. The subject »Technical Thermodynamics« was used as a pilot subject and as an example for other STEM subjects.

The project has been built around the experience of previous e-learning projects, the teaching staff, and students' feedback. With the collected data analysis, didactic and technical specifications were defined to develop the online content with the selected tools: OPAL and ONYX. Many solutions were given to typical issues in STEM subjects such as the use of tolerances, the consideration of consequential errors, etc.

A methodology for a year-to-year improvement cycle has been introduced as a tool to improve the quality of the e-learning material

2 Acknowledgments: We would like to thank the European Social Funds (ESF) and the European Union for the financial support of the project thermoE^{int}, project number: 100251113.

with the existing tools of OPAL/ONYX and the participation of the students. Statistical data is essential to progressively achieve a better teaching material quality.

All planned strategies were successfully implemented and they were generally well accepted by the students as the feedback survey suggests. The results of the tests in the group of the subject »Energy and Thermodynamics« were much improved while the mixed group of »Technical Thermodynamics« shows a slight improvement in the results while the difficulty of the tests were not perceived as easy.

The use of such e-assessment offer helps the student to focus on the key aspects of the subject during their self-study time. As it is done in a voluntary basis, the students can decide if taking the classical approach of the subject or to combine the experience with other modern and technologically more advanced material. It is a way to adapt the subject to individual students, especially the ones with language problems, without having to change the current lecture structure. This is very useful on lectures with a very big audience, where it is more difficult to tackle individual issues. The collected feedback can also help to detect potential teaching concerns and provide information for possible future changes in the lecture content.

Finally, the interaction docent staff-student is still crucial for the good development of e-learning content. The students can also give good suggestions and they can detect errors not seen in the development. This, in conjunction with the statistical analysis makes a very powerful tool for the creation of electronic teaching material and its extension to other potential STEM subjects. The next step in the Chair of Technical Thermodynamics is the continuation of the development of more adaptive e-assessments in the project AdaALF (TU Dresden, Chair of Technical Thermodynamics, 2019) in cooperation with the Hochschule Zwickau and the HTWK Leipzig as a way to incorporate a more flexible system at different universities and subjects. The TU Dresden offers, additionally, a wide support of new technical developments to help to the study success in the current and following years, whose details and updates can be tracked at the study success projects webpage (TU Dresden, 2019).

References

Averell, L. & Heathcote, A. (2011). The form of the forgetting curve and the fate of memories. *Journal of Mathematical Psychology*, *55* (1), 25–35.

Berthold, C., Jorzik, B. & Meyer-Guckel, V. (2015). *Handbuch Studienerfolg: Strategien und Maßnahmen: Wie Hochschulen Studierende erfolgreich zum Abschluss führen* (Stifterverband – Verwaltungsgesellschaft für Wissenschaftspflege GmbH). Essen: SeitenPlan GmbH Corporate Publishing.

BPS Systems (2016a). *BPS System – ONYX Aufgabentypen*. Retrieved from https://www.bps-system.de/help/display/ONYX/Aufgabentypen [09.12.2019].

BPS Systems (2016b). *BPS Systems – Benutzerhandbuch*. Retrieved from https://www.bps-system.de/help/display/ONYX/Hilfe [09.12.2019].

BPS Systems (2019a). *BPS Systems – ONYX Test Suite*. Retrieved from https://www.bps-system.de/cms/produkte/onyx-testsuite/ [09.12.2019].

BPS Systems (2019b). *BPS Systems – OPAL Lernmanagement*. Retrieved from https://www.bps-system.de/cms/produkte/opal-lernmanagement/ [09.12.2019].

Breitkopf, C. (2015). *»thermoE und thermoSA« – Erfahrungen in der Umsetzung von E-Assessment in der Vorlesung Technische Thermodynamik im WS 2013/14 und WS 2014/15*. Presented at the E-Teaching-Day, TU Dresden. Retrieved from https://tu-dresden.de/ing/maschinenwesen/iet/tt/ressourcen/dateien/thermoE/projekt-ergebnisse/thermoSA_etd_2015 [09.12.2019].

Breitkopf, C., Grau Turuelo, C. G. & Banos García, O. (2017). thermoE^int: building e-assessment content for the integration and success of international students in STEM fields. *EUNIS 2017 Congress Proceedings*, 8. Münster.

Breitkopf, C., Köhler, T. & Kretzschmar, J. (2014). *Entwicklung eines kompetenzorientierten E-Assessment für das Fach Technische Thermodynamik – thermoE* (pp. 159–184). Dresden: Arbeitskreis E-Learning der Landesrektorenkonferenz Sachsen.

Ebbinghaus, H., Ruger, H. A. & Bussenius, C. E. (1913). *Memory: A Contribution to Experimental Psychology*. Teachers College, Columbia University.

Freudenreich, R., Grau, C., Breitkopf, C. & Kretzschmar, H.-J. (2017). Thermo-E – Ein Verfahren zur Erstellung elektronischer Übungsaufgaben im MINT-Bereich. *Wege Zu Technischer Bildung*, 177–182. Ilmenau.

Freudenreich, R. Kretzschmar, H. J. & Breitkopf, C. (2016). E-Assess-MINT – Elektronische Übungen im MINT-Bereich. *14. Workshop on e-Learning*, 49–58. Görlitz.

Freudenreich, R. & Lorenz, T. (2015). SPAts – kompetenzorientiertes Online-Self-/Peer-Assessment zur Selbstgesteuerten und Kooperativen Prüfungsvorbereitung. *HDS Jounal 2/2015 Lehrpraxis im Transfer – Hochschuldidaktisches Zentrum Sachsen (HDS)*, 29–32.

Mulop, N., Yusof, K. M. & Tasir, Z. (2012). A Review on Enhancing the Teaching and Learning of Thermodynamics. *Procedia – Social and Behavioral Sciences*, *56*, 703–712.

Murre, J. M. J. & Dros, J. (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *PLOS ONE*, *10* (7). https://doi.org/10.1371/journal.pone.0120644.

Rech, J. (2013, July). *Studienerfolg ausländischer Studierender*. Presented at the DAAD-Wissenschaftswerkstatt, Saarbrücken.

SMWK (2011). *Der Sächsische Hochschulentwicklungsplan bis 2020*. Retrieved from https://www.htw-dresden.de/fileadmin/userfiles/htw/docs/Personalrat/Aktuelle_Informationen/Hochschulentwicklungsplan.pdf [09.12.2019].

TU Dresden (2016). *Statistische Jahresbericht 2015*. Retrieved from https://tu-dresden.de/tu-dresden/profil/ressourcen/dateien/statjb/StatJB2015.pdf?lang=en [09.12.2019].

TU Dresden (2017). *TU Dresden connected to the world*. Retrieved from https://tu-dresden.de/studium/im-studium/ressourcen/dateien/akademisches-auslandsamt/infomaterialien/druck_broschuere_aaa_en_2017.pdf?lang=de [09.12.2019].

TU Dresden (2019). *Überblick Studienerfolgsprojekte – TU Dresden*. Retrieved from https://tu-dresden.de/studium/im-studium/studienerfolg/ueberblick-studienerfolgs-projekte [09.12.2019].

TU Dresden, Chair of Technical Thermodynamics (2019). *AdaALF: Adaptive Aufgabenerstellung, -Lösungsunterstützung und Feedback*. Retrieved from https://tu-dresden.de/ing/maschinenwesen/iet/tt/studium/e-thermodynamik/folgeprojekte [09.12.2019].