

# Best of Two Worlds: Using Two Assessment Tools in One Course

Raoul Deuss<sup>1</sup>, Christine Lippens<sup>2</sup>, Michael Striewe<sup>3</sup>

<sup>1</sup> Antwerp University, Campus Middelheim, Middelheimlaan 1, 2020 Antwerpen, België  
raoul.deuss@uantwerpen.be

<sup>2</sup> Antwerp University, Stadscampus, Prinsstraat 13, 2000 Antwerpen, België  
christine.lippens@uantwerpen.be

<sup>3</sup> University of Duisburg-Essen, Gerlingstraße 16, 45127 Essen, Germany  
michael.striewe@paluno.uni-due.de

**Abstract:** This paper reports on practical experiences with the two e-assessment tools AlephQ and JACK, explains their key features and sketches usage scenarios from two different universities. Using a lecture in accountancy as a concrete example, the paper then presents a successful concept for improving a lecture by introducing both e-assessment systems. Conclusions are drawn on how to improve a lecture by selecting and combining the most suitable features from different tools.

**Keywords:** Practice report, Item types, Usage scenarios

## 1 Introduction

Using e-assessment systems for formative and summative assessments can improve lectures in many different ways starting from reduced effort for grading exams (e.g. [1]) up to improved learning outcomes by more timely and sophisticated feedback (see e.g. [2], [3]). Since there may be many different requirements on what is needed to improve an actual lecture or course, choosing the right e-assessment system is not an easy thing. Some desired features may be not available within a single system or even not available at all. Using a lecture in accountancy as a concrete example, this paper presents a successful concept that achieved to improve a lecture by introducing two e-assessment systems in two steps. In a first step, the lecture was improved by developing and using an own e-assessment system. In a second step, another system from a different university was added to overcome some known shortcomings.

This paper is organized as follows: Chapter 2 presents the e-assessment system AlephQ, its usage at Antwerp University and its key features. Chapter 0 does the same for the e-assessment system JACK from the University of Duisburg-Essen. Chapter 4 discusses the sample case of a lecture in accountancy from Antwerp University that was successfully improved by both tools. Chapter 4.7 presents the lessons learned and conclusions from this case study.

## 2 The e-assessment system AlephQ

The e-assessment tools AlephQ is developed and used at Antwerp University. AlephQ is a solution for some current limitations of question type based e-assessment tools like Questionmark Perception, Respondus and e-assessment tools in Blackboard, Brightspace, Canvas or Moodle. AlephQ offers the teacher a set of building blocks to implement its own questions, with ordering and grading schemes, within a user friendly interface.

This section will focus on the main principles of the software AlephQ, the usage within Antwerp University, the benefits for teachers and students. The section ends with some practical examples to demonstrate the core features of AlephQ.

### 2.1 The main principle: AlephQ as a user friendly building block based system

The set of AlephQ's building blocks consists of "pages", "rows" within these pages, and "elements" in these rows. Two supplementary blocks are "clusters" and "rules". "Versions" are the final building blocks.

Pages – rows – elements:

To build a question, the teacher generates empty rows on a page. Subsequently, the teacher creates on each row the desired number of elements. For each element, the presentation can be selected: read-only text, input – short text, input – essay, input – numeric, check box, drop down, LaTeX, image ... Elements can be organized on rows in a table-structure. On each input-element a grading scheme can be defined at any time during the process of building or grading.

Clusters – rules:

AlephQ can allow students to enter their answers in different but equally correct sequence of elements or rows (for example when asked for an enumeration of symptoms, features, ...) Hereto these interchangeable elements or rows are grouped in a cluster by the teacher during the building phase and one of the 4 sorting and grading algorithms (rules) is selected upon this cluster.

Moreover, clusters can also be used to define a mandatory minimal score to obtain by students on a set of elements.

Versions:

When a whole page is completed by a student in an unexpected, deviating but coherent way, AlephQ can allow a supplementary version of the page as a whole with a deviating grading scheme.

With these building blocks, AlephQ provides the teacher better possibilities to implement their ideal questions in respect to the format, the content and the grading algorithm. AlephQ intends to stimulate self-sufficient users by handy to use functionalities

such as ‘copy and paste’, and Word-compatibility. Mastering how to build an exam autonomously in AlephQ takes an average teacher 3 hours of instruction.

## 2.2 Usage and benefits of AlephQ within the Antwerp University

In de second term of 2014-2015 AlephQ was tested on a few pilot courses. The roll out of the software AlephQ started in the academic year 2015-2016 and turned out to be very successful (cf. **Table 1**).

In the current academic year 2017-2018 AlephQ is used in 92 courses in 9 different faculties of the Antwerp University. 75% of the different electronic exam copies are AlephQ files, whereas 25% of the electronic copies are BlackBoard\* online tests or assignments.

**Table 1.** Number of electronic exams and used application at the Antwerp University

	Total	AlephQ		Blackboard*	
2015-2016	20594	12780	62%	7814	38%
2016-2017	24137	15657	65%	8480	35%
2017-2018	31922	23884	75%	8038	25%

\* BlackBoard examination consists of online tests and assignments

The usage of AlephQ has the following benefits for the teachers at the Antwerp University:

- The time teachers spend on grading exams has reduced significantly (to a factor 10). An example is the course Accountancy where the manual grading of +/-750 exam copies is reduced from three weeks to two days. In most cases several days of manual grading is reduced to several hours.
- More complex exams can be build: the classical criteria of ‘easy to mark’ is no longer valid. Questions that focus on a deeper understanding rather than on reproduction or on multiple choice are within reach with AlephQ.
- Several teachers reinvested their gained grading effort in mid-term AlephQ assessments and home assignments to offer more insight to their students. These AlephQ assessments and assignments are considered as a worthy preparation for the exams and notwithstanding their complexity, are not considered as time consuming for the teacher.

Students also benefit from the software AlephQ: they agree that the AlephQ home assignments are a good preparation for the exams. They are also positive about the detailed personal feedback that AlephQ provides. Even though students may experience a higher level of difficulty (92 courses have changed their exam content when transferring to AlephQ), but the numbers of students that pass their exams do not deviate significantly from the previous exam period.

### 2.3 AlephQ: practical examples

Question 4	
By heating ammonium nitrate decomposes explosively as	
$\text{NH}_4\text{NO}_3(\text{s}) \rightarrow \text{N}_2(\text{g}) + \text{H}_2\text{O}(\text{g}) + \text{O}_2(\text{g})$ (unbalanced equation)	
Question 4a	
Specify the numbers of the particles in the balanced equation. (1 point)	
<input type="text" value="2"/>	$\text{NH}_4\text{NO}_3(\text{s})$
<input type="text" value="2"/>	$\text{N}_2(\text{g})$
<input type="text" value="4"/>	$\text{H}_2\text{O}(\text{g})$
<input type="text" value="1"/>	$\text{O}_2(\text{g})$
Question 4b	
How many liters of gas is formed in total at 550 ° C and 1.15 atm at the detonation of 450 g of ammonium nitrate ( $R = 0,082058 \text{ L}\cdot\text{atm}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$ )? (1 point)	
<input type="text" value="1,16E3"/>	L
Question 4c	
What is the partial pressure of nitrogen gas in the mixture? (1 point)	
<input type="text" value="0,329"/>	atm

**Fig. 1.** Clustering numeric elements – Grading

**Fig. 1** shows a question taken out of the course of chemistry in the science faculty. In question 4a the teachers wishes to assign a score only if all of the 4 answers by the student (i.e. 2, 2, 4, 1) are correct. Therefore the 4 fields of question 4a are grouped in a cluster and a proper grading rule is defined accordingly.

The answer of question 4b relates to the answers of question 4a by the formula:  $1,16\text{E}3 = 5,62 \times (1 + 4 / 2 + 1 / 2) \times 0,082058 \times 823 / 1,15$ . When grading by hand, most teachers would assign a partial mark for a correct use of the formula, even if the absolute value of the student's answer in 4b is incorrect, due to a student's mistake in 4a. AlephQ offers this partial grading functionality: the teacher defines for answer field 4b the dependency and the appropriate partial mark as a marking rule. Supplementary, AlephQ allows the use of ranges on numbers, alternative answers and alternative dependencies.

Vocabulario	
<i>Traduce la palabra entre paréntesis al español. Pon la forma adecuada (por ej. verbo conjugado, singular/plural, con/sin artículo).</i>	
¿Pueden pronunciar ustedes esta frase ( <b>zonder zich te vergissen</b> )	<input type="text" value="sin equivocarse"/> ninguna vez? Yo no puedo.
Necesitamos tener más presencia en ( <b>de sociale netwerken</b> )	<input type="text" value="las redes sociales"/> , como Facebook y twitter.
"Atentamente" es una fórmula de ( <b>afscheidsformule</b> )	<input type="text" value="despedida"/> en las cartas formales.
Nuestra organización tiene muchos ( <b>leden</b> )	<input type="text" value="miembros"/> en el extranjero.

Verzamelde tekst uit Vocabulario					
Uniek Antwoord	Aantal	Prullenbak	Equivalentie	Max Score	Score
	21	<input checked="" type="checkbox"/>	<input type="radio"/> miembros		
cento	1	<input type="checkbox"/>	<input type="radio"/> miembros		
los miembros	1	<input type="checkbox"/>	<input checked="" type="radio"/> miembros	1	1
membres	4	<input type="checkbox"/>	<input type="radio"/> miembros		
membros	9	<input type="checkbox"/>	<input type="radio"/> miembros		
miebres	1	<input type="checkbox"/>	<input type="radio"/> miembros		
miebros	2	<input type="checkbox"/>	<input type="radio"/> miembros		
miembres	14	<input type="checkbox"/>	<input type="radio"/> miembros		
miembro	1	<input type="checkbox"/>	<input checked="" type="radio"/> miembros	1	1
<b>miembros</b>	<b>104</b>	<input type="checkbox"/>	<input type="radio"/> miembros	1	1
miémbros	3	<input type="checkbox"/>	<input checked="" type="radio"/> miembros	1	0,5
miempos	1	<input type="checkbox"/>	<input type="radio"/> miembros		
participantes	3	<input type="checkbox"/>	<input type="radio"/> miembros		
participientos	1	<input type="checkbox"/>	<input type="radio"/> miembros		

**Fig. 2.** Collection of unique given answers for the text element “miembros”

To reduce the time teachers spend on correcting exams, AlephQ creates for each text element a sorted list of unique given answers during the grading process. In **Fig 2** the unique list for the text element “miembros” from an exam of Spanish is shown. The frequency of each answer is indicated. For each text element in the exam teachers can go through the corresponding list and add partial marks for partially correct or alternative answers.

Lists can be compressed as appropriate by leaving out of account white space, punctuation or case sensitivity during the aggregation of the list.

For the average teacher at Antwerp University, the necessary grading time for text elements is reduced by 70% (Q1 = 65%; Q3 = 83%; Average = 72%; Median = 74%)

Casus Delphine Boël	
Delphine Boël is in 1986 geboren uit het huwelijk van Jacques Boël en Sybille de Séllys Longchamp. Zij wordt door hen samen opgevoed. In 1999 verschijnt een boek waarin wordt beweerd dat Delphine eigenlijk een biologische dochter is van Albert van Saksen-Coburg. Delphine bevestigt in de media dat zij dit al sinds haar 12 jaar weet. In 2013 stelt ze de nodige procedures in.	
<b>Vraag 1</b>	
Eerst en vooral moet Delphine een vordering instelling tot betwisting van het vaderschap van Jacques Boël. Geef 2 gronden op basis van de welke die vordering volgens de wettekst niet-ontvankelijk moet worden verklaard? Geef telkens de wettelijke grondslag.	
<b>Grond 1:</b>	bezit van staat - Artikel: 318, § 1 BW
<b>Grond 2:</b>	verval een jaar na de ontdekking - Artikel: 318, § 2, eerste lid BW

**Fig. 3.** Clustering interchangeable rows

**Fig. 3** shows a question taken out of the course of Family law from the faculty of Law. In this example the teacher asks the student to summarize 2 legal grounds (“Grond 1” and “Grond 2”) together with the corresponding law number (“Artikel”) for a given legal situation.

The student is allowed to choose the order in which the 2 grounds are entered. Synonyms or writing errors are permitted. The student is not allowed to combine ground 1 with the law number of ground 2 or vice versa. The student is not allowed to enter the same ground twice.

AlephQ is able to semi-automatically grade this question. In the building process, the teacher groups the two rows with the answer fields in an interchangeable cluster and defines a rule to penalize for duplicate grounds (none or only one ground will be graded). During the grading process synonyms or writing errors can be selected from the list of unique answers (cf. Example 2).

If during the grading process a third ground should emerge to be applicable, this supplementary answer can be added to the grading scheme in back-up rows (reserve rows) as shown in **Fig 4**. A third ground can’t be taken into account for grading by a mere synonym for grounds 1 or 2. As a result, a correct answer will consist of 2 out of the 3 possible grounds.

Reserve draden	
Grond 3:	<input type="text"/> - Artikel: <input type="text"/>

**Fig. 4.** A supplementary answer on a reserve row

### 3 The e-assessment system JACK

At the University of Duisburg-Essen, the e-assessment system JACK is used. It offers simple exercise types like multiple choice or multiple response questions and cloze texts as well as complex exercise types like code submission for programming exercises. In this paper, we focus on the simple exercise types. An important feature for the simple exercise types are JACK’s capabilities to use placeholders within exercise texts that JACK can fill dynamically and individually for each student [4]. It is possible to use placeholders for simple elements like numbers and texts, but also for complex elements like images. Content can be assigned to these placeholders by random selection, by calculations via computer algebra systems (CAS) or via an image generation mechanism (such as plotting a graph via a CAS). Another core feature for simple exercise types is the possibility to split an exercise into several stages that have to be solved one after another. The sequence of stages does not have to be linear, but the choice of the next stage may depend on the input in previous stages or on random selection [5].

The following sections provide an overview on the usage scenarios of JACK at the University of Duisburg-Essen (section 3.1), demonstrate and discuss some sample exercises (section 3.2), and discuss benefits for teachers (section 3.3) as well as for students (section 3.4).

### 3.1 Usage Scenarios

The University of Duisburg-Essen uses JACK in various domains of study and in different usage scenarios covering diagnostic, formative and summative assessments. **Table 2** provides an overview of the size of several scenarios in recent years.

Diagnostic assessments are performed for placements tests for language courses for foreign students. Different tests exist that can be taken either from at home or from a computer pool at the university. Each test can only be taken once. These placement tests solely use C-tests with a single stage per exercise. None of the specific features of JACK regarding dynamic replacement of placeholders are needed in this scenario, as calibrated tests are used which need to be the same for all participants.

Formative assessments are offered to support a wide range of lectures across several institutes of the university, including economics, computer science, mathematics, mechanics, biology, pedagogy, psychology, and language studies. In most of these cases, homework exercises are offered and students may take an unlimited amount of attempts to each of them. Most assessments make use of fill-in exercises or cloze texts with numerical or textual input. Particularly in economics, mathematics, and mechanics many exercises with numerical input are used that rely on dynamic creation of exercise contents via computer algebra systems.

Summative assessments are offered in two different ways: Mid-term assessments every few weeks during the term are used in the institutes of economics and mechanics. In most cases, these assessments are conducted in a computer pool of the university, but there are also some cases of assessments that can be taken from at home. End-term assessments via JACK are used in the institutes of economics and language studies. In summative assessments, the same exercise types and features are used as in formative assessments, with a lower share of exercises using dynamically generated content. In particular, the institute of language studies uses similar exercises for all participants in their summative assessments.

**Table 2.** Numbers of different lectures and courses using JACK at the University of Duisburg-Essen in recent years.

Assessment type	Number of different lectures/courses (approx. sum of participants)			
	Summer term 2016	Winter term 2016/17	Summer term 2017	Winter term 2017/18
Diagnostic	---	5 (570)	3 (1670)	6 (1740)
Formative	7 (1450)	9 (3800)	9 (1770)	10 (2100)
Summative	5 (1230)	5 (1430)	5 (1000)	1 (230)

### 3.2 Sample exercises

This section demonstrates and discusses some system features that exercise authors commonly use in various lectures and domains of study. The demonstration of each feature include a sample exercise from an actual course.

The first sample exercise is shown in **Fig. 5**. It is taken from a first-year course in the institute of economics. The exercise item presents a small table with some randomly generated integer numbers and asks students to enter the result for a specific sum operation on that data in an input field. In this case, not only the numbers in the table are selected randomly within certain bounds, but also the various elements of the sum operation. In particular, the sum does not necessarily run from 1 to 5 (as shown in Fig. 1) and may refer to  $x_i$  in one case and  $y_i$  in another case. Consequently, the correct answer also differs for each instance of this item. Hence, a computer-algebra-system is used to evaluate the correctness of the input by performing the same calculation and comparing the result to the input.

The second sample exercise is shown in **Fig. 6**. It is taken from an introductory course to botanics at the institute of biology. The exercise item presents five formal descriptions of blossoms and asks students to tick the one that fits to a particular family of flowers named by their scientific name. As in the previous example, the item content is generated randomly. However, the mechanisms for content generation are different. Obviously, the scientific name of the family of flowers needs to be chosen from a list of names that matches the contents of the introductory course. For each of these names, one correct answer option exists that needs to be listed. The other answer options are generated by replacing one or more small segments of the correct answer by wrong elements. An exercise author can design precise rules on what to replace and how to replace individually for each family of flowers. This way, the answer options can cover typical mistakes. At the same time, wrong answer options are not too easy to identify, as they do not contain segments that are obviously wrong as it could occur with plain random generation.

#### Frage 1

Gegeben sind:

$i$	1	2	3	4	5	6
$x_i$	7	9	7	7	7	1
$y_i$	5	8	9	1	6	2

Berechnen Sie:  $\sum_{i=1}^5 x_i =$

**Fig. 5.** Sample exercise on mathematics with random elements. Both the numbers in the table at the top of the exercise as well as the elements arrange around the sum symbol are selected randomly within certain bounds for each instance of this exercise.

**Frage 1**

Wie lautet die Blütenformel der Amaryllidaceae?

**Antworten:**

- \* [P(5) A3] G(3)
- \* [P(3+3) A3] G3
- \* [P(5) A3+3] G(3)
- \* [P(3+3) A3+3] G(3)
- \* [P(3+3) A3] G3

**Fig. 6.** Sample exercise on botanics with random answer options. The answer options are created randomly based on rules that assure each answer option to be almost similar but not equal to the correct one.

The third sample exercise is shown in **Fig. 7**. It is again taken from a first-year course in the institute of economics and is part of the exercise pool for homework exercises. The exercise item consists of six stages that must be answered one after another. The first stage of the item introduces a small scenario about two farmers and their products. The text also contains some numbers that are crucial for the calculations to be performed throughout the whole item. However, also the names of the products, the names of the farmers and even their gender are generated randomly to make the exercise more vivid. All random choices are applied to all six stages of the item, but user input may be considered as an additional parameter for content generation later on. Hence, the content of one stage can also depend on input from previous stages, which allows to account for subsequent faults or to allow students to choose one of several appropriate methods for solving a particular problem.

### Frage 1

Die beiden Landwirte **Franzi (F)** und **Margarete (M)** leben und arbeiten in einem kleinen Dorf. Die Bewohner des Dorfes haben eine spezielle Diät für sich entdeckt und ernähren sich seitdem nur von zwei Gütern: **Erdbeeren** und **Käse**. Den beiden Landwirten bleibt also nichts mehr anderes übrig als sich auf die Produktion von diesen beiden Gütern zu spezialisieren.

**Franzi** kann entweder **2 Kilogramm Erdbeeren** oder **6 Kilogramm Käse** pro Stunde produzieren.

**Margarete** kann **5 Kilogramm Erdbeeren** oder **8 Kilogramm Käse** pro Stunde produzieren.

Die möglichen Produktionen pro Stunde sind in folgender Tabelle zusammengefasst:

	Franzi	Margarete
Erdbeeren	2 kg/h	5 kg/h
Käse	6 kg/h	8 kg/h

Als ersten Schritt wollen wir ermitteln, welcher Landwirt jeweils in der Produktion für welches Gut den **absoluten Vorteil** besitzt.

**Kreuzen Sie dazu bitte entsprechend an, welcher Landwirt für welches Gut den absoluten Vorteil besitzt!!**

	Franzi	Margarete
Erdbeeren	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Käse	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Feedback:

Richtig. Gut gemacht. Sie haben alles richtig angekreuzt.

Punkte: 100/100



**Systemnachricht:** Gut! Sie haben die 1. von 6 Fragen absolviert.

### Frage 2

Nun interessieren wir uns dafür, welcher Landwirt für welches Gut einen **komparativen Vorteil** besitzt.

Dazu berechnen wir uns zunächst, welche Opportunitätskosten jeweils bei der Produktion der beiden Güter anfallen.

**Tragen Sie in die Tabelle die bei der Produktion jeweils anfallenden Opportunitätskosten ein!**

	Franzi	Margarete
Opportunitätskosten für 1kg Erdbeeren	<input type="text"/> kg Käse	<input type="text"/> kg Käse
Opportunitätskosten für 1kg Käse	<input type="text"/> kg Erdbeeren	<input type="text"/> kg Erdbeeren

**Fig. 7.** First two steps of a larger sample exercise on economics. Several elements of the exercise including numbers, names and terms are generated randomly. The second step of the exercise ("Frage 2") is only revealed after a student made a submission to the first one.

### 3.3 Benefits for teachers

Items as the ones shown in section 3.2 can be used both for formative and summative assessments. In formative assessments, it is sufficient to define one item template with carefully designed rules for random content generation to produce a virtually unlimited amount of item instances for practice. In summative assessment, this amount of different but similar item instances can help to avoid plagiarism. Consequently, one of the main benefits for teachers and exercise authors is the relatively low amount of time they

need to create a large amount of items. Although it consumes significantly more time to design an item with dynamic content generation than to design a static item this is outweighed by the reduced need to produce several items.

Items with several stages allow to develop large, complex scenarios and thus to create more challenging exercises than exercises with a single stage. At the same time, they allow to ask for intermediate results instead of just asking for final results and offer options to handle subsequent faults and alternative ways of problem solving. Again, this comes to the cost of higher effort for designing a single item.

Particularly in the case of mathematical content, teachers can also rely on the capabilities of computer-algebra-systems for evaluating the correctness of answers instead of figuring out corner-cases by hand. Thus, exercises may also ask students to enter complex terms instead of single numbers and computer-algebra-systems can automatically handle transformations on those terms and determine their properties. In turn, these capabilities can also be used to get fine-grained control over the generation of exercise contents.

### **3.4 Benefits for students**

Students can profit by the benefits for teachers in several ways. First, they get more opportunities to practice in formative assessments, as they can work on dynamically generated items repeatedly. In summer term 2018, the introductory course to botanics mentioned above used a total amount of 151 exercises. On average, students who made at least 20 submissions at all (and thus showed at least some serious interest in the exercises), made 2.04 submissions per exercise. 15 out of 133 students made more than 3 submissions per exercise. Similar figures could be observed at the institute of economics with an average of 2.76 submissions per exercise in a course on statistics in winter term 2017/18 and an average of 2.09 submission per exercise in a course on microeconomics in summer term 2016.

As a second benefit, student get more opportunities to show their individual competencies, if items with several stages are used. In these cases, it is not necessary to solve the problem completely, but grades and feedback can also be assigned to smaller steps of the solution. Students appreciate that as fair grading particularly in summative assessments.

As a third benefit, students can work on more practically relevant and realistic cases, if items with several stages and vivid content are used. This helps them to not only learn the mechanics of problem solving or calculating, but also to learn how to identify and extract the relevant pieces of information from the question texts.

## **4 The use of the complementary packages AlephQ and JACK**

### **4.1 The use of AlephQ and JACK to activate large groups of students**

In the first term of the coming academic year 2018-2019, we intend to roll out the Jack software package in addition to the software package AlephQ in the Accountancy course of the faculty of Applied Business Economics at Antwerp University.

### **4.2 The Accountancy course**

About 700 students are enrolled in the Accountancy course, which is considered to be one of the more difficult subjects, with a relatively low pass rate on the exam. This is mainly due to the strong differences in students' prior knowledge and adaptation problems some students face in their freshman's year.

Accountancy students need to digest accounting principles and regulation and be able to apply these to practical cases. That is why practicing regularly after classes is essential to make a pass grade on the exam.

Accountancy is taught over two terms in a year. And the end of each term a partial exam is set. In each term we teach six units, consisting of a classroom lecture followed by a practice tutorial in smaller groups. To force student engagement, we set three homework assignments in the first term (each after finishing two units), and just one in the second (after the fourth unit).

We use the classroom lectures to explain both law and accounting principles. In the tutorial sessions, we start with some basic exercises in order to lift all students above the same minimum threshold. We then expect that the students use the compulsory homework assignments to upgrade their knowledge and skill to the required final level, which corresponds to the difficulty level of the exam.

We support our students by giving them access to a study resource of 50 short knowledge video-clips covering all the topics. Each clip is linked to the unit to which the topic belongs. Students also get access to a set of additional solved problems. Finally, we offer them and monitor a Blackboard forum to ask questions about the course material.

### **4.3 Use of AlephQ**

Thanks to AlephQ, we are able to have students solve numerically extensive cases (as shown in **Fig. 8**) and give them personalized feed-back on their answers in less than 24 hours after the submission deadline.

Because the difficulty level of the problems given in the assignment matches that of the problems given on the exam, and because the computerized exam is also set in the AlephQ environment, students can benefit from a familiarization effect: the assignments are an excellent preparation for both the look and feel and the difficulty level of the exam. This set up forces the students to keep up with the subject matter during the year and allows them to adjust their study approach within the semester.

Because of its user-friendliness, AlephQ makes it easier for teaching staff to develop new accounting cases. One should be able to develop a complete, quantitatively extensive question set in one day. Without AlephQ, this would have been feasible too, but it would have taken several weeks of manual marking to correct such an assignment.

**DEEL 10: BANKSALDO**

Bereken de evolutie van het banksaldo, op basis van alle gegevens.

**Regels voor correct invullen:**

Afname van het banksaldo: negatief bedrag invullen  
 Toename van het banksaldo: positief bedrag invullen  
 Nullen expliciet invullen, alle velden invullen

BANKSALDO			
	2013	2014	2015
<i>Beginsaldo bankrekening</i>	50.000,00	66.970,00	29.231,19
Deel 1: Overname-akte	-46.630,00	0,00	0,00
Deel 2: TopStop	0,00	0,00	0,00
Deel 3: Segways SMV	-2.400,00	-2.400,00	-2.400,00
Deel 4: Segways STT	6.000,00	6.000,00	6.000,00
Deel 5: Audiogidsen	0,00	-600,00	-600,00
Deel 6: Bestelwagen	0,00	-6.738,81	-6.738,81
Deel 7: Oefenterrein	0,00	0,00	0,00
Deel 8: Indoor parcours	25.000,00	-69.000,00	-24.000,00
Deel 9: Lessen via 10-beurtenkaarten	35.000,00	35.000,00	35.000,00
<i>Eindsaldo bankrekening</i>	66.970,00	29.231,19	36.492,38

**Fig 8.** Detail of the assignment

The question set is not randomized, i.e. all students solve the same (complicated) case. One advantage of a unique problem case is that students can collaborate in study groups and learn from each other as well as from the lecturing and support staff.

We treat the compulsory assignments in the Accountancy class as a learning opportunity as opposed to an assessment of competence. That is why teaching staff is answering questions about the assignment in the dedicated forum, up to multiple times a day, to help the students on their way. The forum gets consulted about 1000 times per day during the brief period that the assignment is "live". The third assignment got a record 9800 student interactions (see **Fig. 9**).

	Assignment 1	Assignment 2	Assignment 3	Assignment 4
Day 1	443	575	1743	221
Day 2	1030	1251	2749	748
Day 3	986	1663	2541	1039
Day 4	1122	1066	2079	483
Day 5	1977	1449	746	534
Day 6	582	1480		673
Day 7		800		731
Day 8				782
Day 9				1383
<b>Total</b>	6140	8284	9858	6594
<b>Average</b>	1023	1183	1972	733

**Fig 9.** Forum Assignments: Amount of views

Using a unique problem set also allows the lecturer to post a feed back video, in addition to the personalized automated feed back.

Because the assignment is administered in an uncontrolled environment (it is a homework assignment with multiple group interactions), marks on the assignment do not count towards the final exam score. By contrast, students that don't participate or that get very low scores (less than 25%) are penalized by having their final score reduced by 2 points. They can avoid this penalty by submitting additional remedial homework (which is manually set by the teaching staff and which thus inevitably means additional work for them).

#### 4.4 Deploying Jack

We have recently invested the time and effort to create a set of drill-questions for Accountancy in the Jack environment. Jack is an excellent tool to generate a pool of randomized analogous problem sets. These drill-questions will help students to reach the basic skill threshold more quickly through this digitization.

Jack is also capable of handling larger quantitative concepts through its parametrization features. This requires the staff to make a one-time effort to create it, but then Jack can deliver a theoretically limitless amount of practice materials.

#### 4.5 Advantages to the Students

For each unit, Jack exercises are offered in a module of the learning environment as materials for self study. A big plus is that Jack offers the student instantaneous automated feed back. In addition, the offering of training materials is large due to the parametrization.

As such, these quantitatively more extensive Jack exercises form a bridge between the basic skills from the tutorial sessions and the more elaborate problem sets given in the AlephQ environment.

#### 4.6 Advantages to the teaching staff

Although the initial time investment in the Jack set up is non-negligible, the effort is a one-off and the problem sets then become reusable year after year. They can also automate the remedial homeworks that teaching staff now administers to those students that have failed an AlephQ assignment.

In summary, Jack complements our existing study materials and provides an opportunity to automate the remedial work.

#### 4.7 Conclusion

In this paper, we presented the case of the Accountancy course at Antwerp University, improved by rolling out both e-assessment packages AlephQ and JACK. These software packages offer complementary functionalities and complementary benefits which results in a richer learning experience for students than the standard question types from the classical learning-management-system could offer.

The combination of Jack and AlephQ allows the teaching staff to support students to master essential principals, skills and knowledge, necessary to master a discipline and to properly prepare for the exam. By offering quick and detailed feedback on a regular base, students have the opportunity to review and to optimize their study strategy. AlephQ and JACK make this feasible for large groups of students.

## References

1. Ridgway, J.; Mccusker, S.; Pead, D.: Literature Review of E-assessment. A NESTA Futurelab Research report - report 10 (2004).
2. Malmi, L.; Korhonen, A.: Automatic feedback and resubmissions as learning aid. In: IEEE International Conference on Advanced Learning Technologies, 186-190 (2004)
3. Venables, A.; Haywood, L.: Programming students NEED instant feedback! In: Proceedings of the fifth Australasian conference on Computing education - Volume 20, Australian Computer Society, Inc., 267-272 (2003).
4. Schwinning, N.; Striwe, M.; Savija, M.; Goedicke, M.: On Flexible Multiple Choice Questions with Parameters. In: Proceedings of the 14th European Conference on e-Learning (ECEL) (2015).
5. Schwinning, N.; Schypula, M.; Striwe, M.; Goedicke, M.: Concepts and Realisations of Flexible Exercise Design and Feedback Generation in an e-Assessment System for Mathematics. In: Joint Proceedings of the MathUI, OpenMath and ThEdu Workshops and Work in Progress track at CICM, co-located with Conferences on Intelligent Computer Mathematics (2014).