Loughborough University

This item was submitted to Loughborough's Research Repository by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Online assessment for University admission: goals, problems and experience

PLEASE CITE THE PUBLISHED VERSION

PUBLISHER

© Loughborough University

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Turau, Volker, Dietrich Fahrenholtz, and Marcus Venzke. 2019. “Online Assessment for University Admission: Goals, Problems and Experience”. figshare. https://hdl.handle.net/2134/2009.

ONLINE ASSESSMENT FOR UNIVERSITY ADMISSION: GOALS, PROBLEMS, AND EXPERIENCE

Volker Turau, Dietrich Fahrenholtz, and Marcus Venzke

Online Assessment for University Admission: Goals, Problems, and Experience

Volker Turau, Dietrich Fahrenholtz, and Marcus Venzke
Hamburg University of Technology, Department of Telematics
Schwarzenbergstraße 95, 21073 Hamburg Germany
turau,fahrenholtz,venzke@tu-harburg.de

Abstract

This paper reports on the experience gained during the implementation and operation of an online assessment system for recruitment of engineering and computer science students. We identified problems that have to be solved to make online assessment a worthwhile building block of university admission. The presented work includes techniques for evaluating and monitoring online assessment processes.

Introduction

After the establishment of state-controlled selection procedures pertaining to university admission, recent changes in German regulations give universities more choices over whom to select and whom not. Universities favour this kind of autonomy which allows them to select students who closely fit to their courses and educational profile. However, abiding by the procedures incurs a lot of work, especially if the number of applicants is high. Developing an efficient admission for the Hamburg University of Technology also required the consideration that potential students apply from all over the world. Inviting all of them to Hamburg is simply not feasible.

For this reason we have decided to implement the first step of our admission process as *online assessment*. This term originally meant assessing students' performance during their studies with the aid of Web-based technology. We utilize the term with its broader meaning that also includes recruitment of potential students using the same technology from all over the internet. Making online assessment the initial contact for candidates prior to a traditional paper based application, results in several benefits:

- Candidates weakly interested in the course or university are challenged.
- Potential students with low prospects are deterred.
- The university's educational profile is sharpened.
- The number of good candidates is likely to increase.

- Candidate profile statistics can be gathered eventually leading to the development of new courses.

Only few publications cover online assessment for selecting students. Harding reports about experience with two online-tests for university admission in sufficiently secure computer pools and compares them to tests on paper [4]. Other authors apply online assessment in university courses for placing students in parallel classes of different levels [9] and for informing students about their prospect to succeed a class [2].

We have developed a system for online assessing candidates for engineering courses and computer science already in use at the Hamburg University of Technology. This paper reviews our procedures and experience. Sections 2 and 3 summarize the goals and issues of online assessment in general. Our assessment procedure is described in section 4 followed by our technical system in section 5. Section 6 reports our experience, section 7 concludes.

Goals of Online Assessment

The primary goal of not only our university is to select and register those potential students who will successfully finish their studies with high probability. A further goal is to decrease the number of students who leave university without degree. We would also like to identify students' learning prerequisites and whether or not university quality standards will be met in the future. The certainty with which a student decides in favour of a particular course is an essential factor for the success of his1 studies. Good knowledge about specific study conditions and detailed knowledge about the university's technical standards is the foundation of a certain decision [7]. When potential students subject themselves to an assessment, they made sure they know what is being requested and had a good look at the standards. For us it is particularly interesting to find out where good to outstanding candidates have gone to school and whether there is a bias towards one or more schools.

Assessments are classified into two categories: One is *self-selection*, and the other is *external selection*. Self-selection is to guide the potential student to

1. better choose the most appropriate course from among the offered,
2. become aware of course contents to be imparted,
3. get an idea, if he has enough interest in the matter,
4. decide where he stands in terms of his knowledge (e.g., before exams).

Furthermore, self-selection, in form of a self-test, limits the effort universities put into selecting desired students. This is especially true, if applicant numbers are high. On the other hand, external selection, although more costly in terms of time and faculty involved, may lead to more valid results when it comes to register a potential student. Selection interviews, for example, can

[1] The online assessment system is used by female and male candidates, in order to improve the readability of the text only the male form for gender-related pronouns is used in the following.

be used to sort out candidates who lack required abilities (e.g., discipline or intrinsic motivation) that cannot be assessed otherwise.

The main advantage of online assessment is that potential students can take the test when and where they want. This is especially important for potential students who come from abroad. Also, large volumes of potential students can be assessed simultaneously. But that inevitably calls for a high degree of automation, e.g., special software-tools need to be developed that produce reliable assessment results without human intervention. Moreover, online assessment establishes confidence in the measurement of potential student performance and can be used to give a performance trend analysis over the course of time [5].

Indispensable properties of every online assessment are:

1. Each selection criteria can be verified objectively.
2. All selection criteria apply to every potential student.
3. Discrimination of disabled persons should be minimised.
4. Significance of established university entrance qualifications is retained.

Also an important issue is what kind of feedback an online assessment gives to potential students. Will they learn correct answers to all of the questions previously assigned to them, or just a fraction, or none? At any rate, feedback after being assessed should give the candidate an impression of the course rationales and requirements he wants to register with and assess if he has enough interest in the matter.

Issues with Online Assessment

First of all, when deciding to employ an online assessment, its success is highly dependent on the candidate's degree of familiarity with computer and Internet use in particular. Since we are assessing candidates for engineering courses, it is reasonable to assume this requirement.

One of the major obstacles to widespread adoption of online assessment is authenticity and identity of participants. Testing remotely, another person can sit the test for the potential student. Harding solves this issue by having tests in sufficiently secure computer rooms only [4]. Unfortunately this restricts the number of candidates and does not allow taking the test when and where they want. Conducting tests over the internet we can only appeal to candidates to be honest, because when starting a course they will have to work on their own.

It is also conceivable that over time questions and solutions appear on some Web site or in a forum, which means the test will eventually lose its significance. The only way to counter this is to have highly parameterized questions and continuous revisions and additions to the catalogue.

Moreover, there is the notorious threat that any Web site could suffer outages let alone be taken down by a hacker. To prevent this, the usual precautions of Web security must be applied.

From the point of view of comprehensive statistical analysis, it is desirable to gather as much candidates' data as possible. But demanding a large amount of information might, on the other hand, put off potential students and tarnish the university's reputation. Thus, finding an adequate balance, i.e., what data is indispensable for meaningful assessment statistics as opposed to data with which comprehensive cross section analyses could be made requires prudence. At any rate, for meaningful statistics, it is important to filter out users who do the online assessment without the intention to apply for a course. Reasons might be interest in the university, the online assessment procedure, proving oneself capable to pass the self-test, or simply fun. Detecting such "fun users" is a challenging task. Section 6.2 describes our approach how to cope with them.

Although interesting, online assessment is unable to measure student efficacy, stress, motivation, and creativity because these variables are outside pure knowledge assessment. A further interesting concern is the legal aspect of online assessment. Will its results remain valid when contested in front of court?

Our Online Assessment Implementation

We believe that every assessment needs to be "easy" and "effective". We follow the definition of Blandford et al. who define easy as a term that “translates into less work for those carrying out assessment" and an "effective assessment method" is "one which, in the opinion of most faculty, produces results that lead to the improvement of a course, program, or curriculum" [1]. To partake in our online assessment, potential students, first, need to register with our Web site, i.e., they have to enter personal data such as first name, last name, date of birth, gender, home town, name of school, and valid email address. Their personal data is guaranteed to be used for purposes of the online assessment only and stored securely. At the end of registration, they obtain a unique identifier valid for one personalized copy of the self-test and later interactions with the university by email. Using this identifier, a potential student can request his personalized questionnaire, i.e., a fixed number of random questions are drawn uniformly from a pool of many questions. Thus we ensure potential students cannot cheat by filling in answers they sought out from a previous participation under, e.g., fake personal data. Each question is accompanied by an answer of one of three possible types: a) Multiple choice (only one answer is correct), b) Multiple response (no, one, or more answers are correct), and c) free-format text. Every self-test is scheduled to last 60 minutes and potential students need to submit their solution within this time frame.

A candidate need not be online during this time frame. He can solve his self-test while being offline and submit it later when he is done. Solutions arriving

out of time will not be considered. If, for some reason, a potential student can substantiate his not being able to submit his solution on time, he can get a second change and redo the self-test. The questions contained in an online assessment are taken from several categories, e.g., reading comprehension German, reading comprehension English, numeracy, general knowledge, etc. Categories have different weights in terms of number of questions per category contained in a questionnaire. Shortly after submitting the solutions, potential students receive a PDF document by email consisting of their questionnaire with their answers and a ranking figure. However, they are not advised which questions were answered correctly and which not, so they only learn if they are within the top 10, 20, 30, … percentile of all potential students who partook in the self-test so far.

The number of correctly answered questions is not an automatic criterion for the admission decision; the test result enters the selection process in form of an essay. In this essay applicants have to discuss how they mastered the test and their personal preliminary estimate. Candidates have to write this essay as final step of their application. The discussion shall also show the understanding of their strengths and weaknesses with respect to the favoured course. In addition candidates shall explain their decision for the course and tell about unrelated activities such as sport, hobbies, social commitment, etc. This gives assessors an embracing picture of the candidate.

Finding suitable questions to be used in personalized questionnaires is not a trivial task. The quality of the questions determines their informational value and is thus vital for the overall success of online assessment.

Questions need to test knowledge from different disciplines and the levels of difficulty should vary eventually leading to a clear distinction between applicant skills. Our personalized questionnaires are not meant to be fit to just one course but to all engineering programs. Thus the spectrum of questions needs to cover the engineering profession in general with emphasis on mathematics. We appointed representatives of professors to jointly identify well-chosen questions with varying levels of difficulty from all disciplines that potential students have to answer. Finally, technical support reviewed these questions to see if they fit to the online test system's model and if so transformed them into the required format.

The System

Systems for online assessment exist commercially or can be implemented for a specific university. We have analysed commercial systems. These provide many features but none met the requirements from section 1, fitted to the university’s general conditions and could be integrated into the systems used by the administration. We thus implemented a new system using freely available software such as the Apache Web server and the MySQL database system. Figure 1 depicts the relationships between participating users and system components. The application software is implemented in Perl and PHP and runs on standard PC hardware.

To avoid discrimination of potential students due to their hardware or software preferences, browser independence is a distinct feature of our online assessment system and the code was proven to run on the majority of Web browsers. The only feature we implemented in JavaScript to show in the browser is a clock that displays the remaining time to solve the test. The Web interface and questionnaires can be configured to adapt to different languages to allow for foreign candidates.

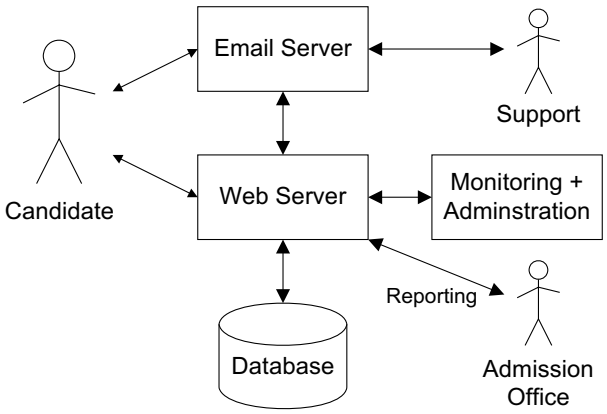


Figure 1. Environment of the online assessment procedure

A general issue with questions containing mathematical content is how to render complex formulas. We see four representation options:

a) bitmap graphics
b) special HTML mark-up (e.g., square root sign)
c) special mathematics mark-up (MathML)
d) dedicated graphics container (e.g. plug-in or Java applet)

We excluded options c) and d) for interoperability reasons, which are lack of required features, non-availability of plug-ins, and prevention of download for security reasons. Special HTML mark-up would be an elegant solution, but HTML does not support all mathematical symbols and decreases the readability of formulas, misleading potential students into making mistakes. Despite all disadvantages of bitmap graphics (being static and not scalable), we decided to use this option that ensures at least readability and interoperability.

Our online assessment system supports import and export of questions as specified by the IMS question and test interoperability specification [6].

System Evaluation

Even a semi-automatic selection of potential students is a delicate task, especially if the number of applicants is only marginally higher than the maximum number of admissions. A rejection of a highly talented applicant should not happen. It is therefore necessary to carefully test and evaluate the

complete procedure prior to deploying it and during its execution. The following aspects should be covered:

- Software
- Overall principle of procedure
- User acceptance
- Suitability and quality of questions

For testing our software, well known techniques from software engineering can be applied, while for evaluating the quality of questions there are only a few techniques available [3]. We propose the adoption of the following means:

- Evaluation with sample groups
- Monitoring the application in operation
- Analysis of ex post data

1.1 Sample groups

The proposed questions were evaluated prior to their usage in the online assessment. Different groups of test persons took the test including professors, post graduate students, and second term students. Even though advisable, it was difficult to include persons from the target group. While performing their sample tests, persons were observed and asked for their opinion while filling in their questionnaire. In addition, analyses of their results led to an indication if the levels of difficulty are suitable. Apart from software testing, following aspects have been evaluated:

- Overall principle of the assessment
- Comprehensibility of the introducing instructions
- Time allowed for the assessment

Eventually this phase led to the final set up comprising of the catalogue of questions with which the real online assessment was started.

1.2 Monitoring

Monitoring the application in operation is accomplished by selected tools. Apart from general ones such as log file analyzers, dedicated tools are needed. The design of such tools should be an integral part of the complete online assessment development process.

The following list contains criteria to evaluate the suitability of individual questions and their compilation for a complete test:

- *Questions with a significant high number of wrong answers.* Possible reasons: ambiguous or incomprehensible formulations (e.g., double negation, unclear language), provided solution is wrong, too difficult.
- *Questions with a very high number of correct answers.* Possible reasons: too suggestive formulation, too easy.

- *Questions that were not answered in a significant number of cases.* Possible reasons: incomprehensible formulation, too difficult.
- *Skewed distribution of overall grading.* Possible reasons: overall level of questions too high or too low, unbalanced mixture of severity of questions in a category.

Figure 2 depicts the distribution of correct answers over a period of 6 months. Marking of the tests was based on rigid scoring, i.e., marks were only awarded for a fully correct answer and no negative scoring was applied. The chart shows that the test achieves a candidate distribution over the possible spectrum. More than 45% of the candidates answered more than 50% of the questions correctly. Brown et al.[3] as well as McAlpine[8] present statistical measures that help identifying the usefulness of individual questions. They aid in finding out whether a question accounts for discrimination of stronger to weaker applicants. This is done by looking at the correlation between the performance of a sub-group of applicants on a particular question and their overall scores in the assessment. Facility is a simple measurement to determine the difficulty of a question by dividing the number of correct responses by the total number of responses. Both measures can easily be integrated into a monitoring tool.

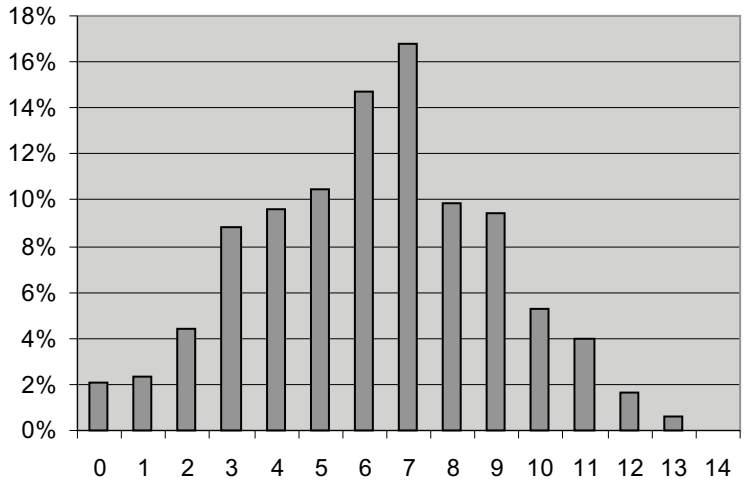


Figure 2. Distribution of correctly answered questions (out of 14 questions)

As a consequence of monitoring, some questions and their corresponding answer patterns should be carefully analyzed. This process leads to the revelation of ill-posed questions or incorrect answers so that those questions can be revised or even discarded.

It has been suggested that the time allowed for an objective test should not exceed 90 minutes. Experiences gained with the sample groups, let us fix the test duration to 60 minutes completion time. To evaluate the appropriateness of this duration, the following quantities are helpful:

- Number of questions answered in time
- Number of late submissions

Figure 3 depicts the distribution of submission times. We interpreted the figure in such a way, that the allowed time span was respected by the majority of participants. We also found almost 70% of all valid applicants answered all questions, an indicator that the number of questions fits to the time span.

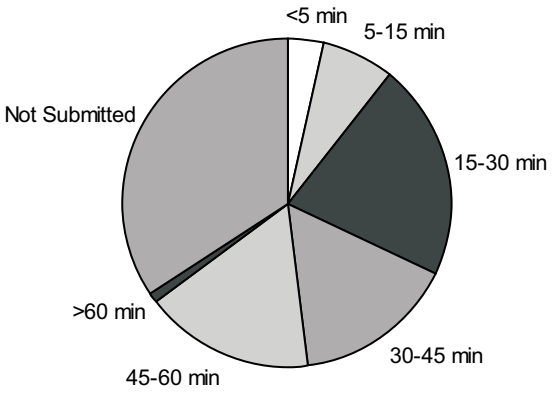


Figure 3. Distribution of submission times

For meaningful statistics, it is important to filter out those users who sit the test without intending to apply for university. Although detecting users who do the test seriously is extremely difficult. Even a missing paper application is not a valid indicator for a “fun user”. It could be, for example, that a serious applicant might have changed his mind after participation in the test. On the other hand, not serious users can be detected, if they are sitting the test in an unexpected way. For example, they might provide senseless names (e.g., "X Y" or "Mickey Mouse") or submit answers too fast not even being able to read the questions properly. Almost one third of the applicants who requested the test did not submit it. This third should not be included in the monitoring process and filtered out. We found the following criteria useful for this filtering step.

- No submission of completed test
- Completion time less than 10% of available time
- Personal data such as name or date of birth is meaningless
- Only a few number of questions are answered

There is conflicting evidence about the extent to which the sequence of questions influences applicant performance of solving the test. Monitoring the distribution of correct answers with respect to question sequence should therefore be performed and appropriate modification should be carried out.

Since the online assessment is a Web-based application the usual precautions have to be installed. This includes a careful analysis of log files to detect attempts of misuse or even attacks, for example.

1.3 Ex post data

The objectives of our endeavour were, first, to support students in their decision whether they are qualified for a course and, second, to support the selection of qualified students by the university. To verify whether these goals have been reached, several possibilities were available:

- Interview applicants at a later time about the impact of their assessment result on their decision process.
- Make the test with second year students and compare their results with those from applying ones.
- Redo the test with students who are studying at the university for one year and compare their test results with their examination results (e.g., grades, overall progress in course)

Conclusions

Designing an online assessment system for university admission requires some effort for finding suitable procedures and questions. However, with careful planning and systematic monitoring it is a viable alternative, if assessment needs to be accomplished with limited personal and financial resources. Our system has been tested as optional part of applications for the winter term 04/05 and is now obligatory for the ongoing and future application periods.

Acknowledgments

We thank R. Jung for implementing the core of the system, many colleagues for discussing about the questions, and the volunteers for taking the test. This work was partially supported by the Multimedia Kontor Hamburg in the context of the program *E-Learning und Multimedia in der Hochschullehre*.

References

[1] D. Blandford and D. Hwang. Five Easy but Effective Assessment Methods. Proc. of the 2003 SIGCSE, pp. 41-44, 2003.

[2] G. D. Boetticher, et al. Using a Pre-Assessment Exam to Construct an Effective Concept-Based Genetic Program for Predicting Course Success. Proc. of the 2005 SIGCSE, pp. 500-504, 2005.

[3] G. Brown, J. Bull, and M. Pendlebury. Assessing student learning in higher education. Routledge, 1997.

[4] R. Harding. Thinking Skills Tests for University Admission. Proc. of the 8th CAA Conference, 2004.

[5] K. D. Hopkins. Educational and Psychological Measurement and Evaluation. 8th ed., Allyn & Bacon, pp. 1-15, 2004.

[6] IMS. IMS Question & Test Interoperability Specification. IMS Global Learning Consortium, June 2004. Version 2.0.

[7] I. Lischka. Auswahl der Studierenden durch die Hochschulen - ist nun blinder Aktionismus angesagt? Zeitschrift für Hochschulentwicklung und -politik, vol. 4: pp. 144-150, 2004.

[8] M. McAlpine. Principles of Assessment. Bluepaper No. 1, CAA Centre, University of Luton, UK, February 2004.

[9] L. K. Soh, et al. Designing, Implementing, and Analyzing a Placement Test for Introductory CS Courses. Proc. of the 2005 SIGCSE, pp. 505-509, 2005.