

The Review-System

A cooperative learning environment

Dr. Jochen Koubek

Humboldt-Universität zu Berlin

Institut für Informatik

Unter den Linden 6

10099 Berlin

Germany

jochen.koubek@hu-berlin.de

Abstract

Study will remain a social process even in times of electronic media. Especially the examination and correction of exercises by a teacher helps shaping individual learning-paths. But teachers are cost-intensive. With large learning-groups more often than not procedures of machine-based corrections are applied, providing poor feedback to the students.

The *Review-System*, developed at the Humboldt-Universität zu Berlin, is a web based learn-management-system, whereby students mutually peer-review their exercises. Every student receives an anonymized random selection of works form other students for correction and review. The challenge lies in finding incentives to assure detailed and significant rapports. In certain learning-situations the system encourages a high engagement of the students due to the social dynamics. Additionally the organization of virtual learning communities without a central corrector can be interesting not only in times of notoriously low education-budgets.

1. Initial Position

The course *Digital Media* consisted of four hours lecture and two hours practical course. The lecture treated subjects like the historical background, technical functions and systematic of digital media. The practical course followed a hands-on approach, considering that media competences are primarily based on practical skills. Concrete design exercises taught the basics of media production. They were treated with standard software, mostly Adobe's Creative Suite, i.e. Indesign, Illustrator, Photoshop and GoLive. The exercises aimed at working on basic techniques of media design. Each exercise has been clearly defined with formal requirements to be fulfilled. The realization of these requirements with the corresponding software

tools was explained during the practical course which turned out to be a kind of tutorial for the actual software. Students were supposed to be present during the two hours but could be absent if they already knew how to work with the specific tool.

The content of the exercise could be freely chosen by the students. Working on proper projects was therefore explicitly encouraged. This was a highly motivating factor, since many students worked on projects they found useful in some way.

During the semester 2003/2004 more than 90 students took part in the course *Digital Media*, an amount evoking some serious problems:

- During examination: 90 oral examinations at 30 minutes each plus time to discuss the results summed up to nearly two weeks non-stop examination taking more time than the proper course. This problem is still unsolved, since the regulations assure the students a right of individual oral examinations.
 - The solution to reduce the number of participants is unsatisfactory for students should have the possibility to follow the courses they like.
- During practical course: 90 students faced one tutor, the author of this article, who had to choose between two unpleasant alternatives:
 - Sacrificing most of his time to the correction of exercises. Reviewing a document and writing a short feedback takes about 10 minutes. 90 documents would approximately demand 900 minutes or 15 hours. Apart from the fact that reviewing masses of media documents is a rather dull task, such an effort would have consumed a disproportional part of the working hours.
 - Diminishing the feedback. If you cut the written review, the correction time for a single document can be dramatically diminished. But in this case the students get virtually no constructive feedback but a simple score. No good learning results can be expected on this way.

Reducing the number of participants was not possible once the semester has started as well as not desired for the reason mentioned above.

Given the choice between two bad alternatives one is well advised not to choose any of them. To solve this dilemma the review-system was developed. It is based on a peer review approach where each document is evaluated by several experts. To the students it has been presented as a game with fixed rules implemented in a web-based platform.

2. The Rules

Inscription

The participation in the review-system required a one-time inscription. Every student wishing to work on one or several exercises had to fill in a form with his name, a registration number – usually the student number – and his email-address. The link between registration number and email-address was necessary to distribute the review-tasks. The additional registration of the name is necessary if the results are to be assigned to concrete students. If the review-system is used in an anonymous environment as described below, the inscription of the name is not necessary.

Structure of an exercise

Every two weeks an exercise was presented, demanding the realization of a document. Two weeks were considered for the preparation of the corresponding document. The exercise description included:

- The date of delivery. Documents submitted after this date could not be taken into account.
- The file format and therefore the file-extension of the document, i.e. PDF, PSD, MOV etc.
- A name convention for the document. This was always *[registration-number].[File-Extension]*, e.g. *1234567.PDF*.
- The formal requirements of the design. The reviews had to be based on these requirements.
- Propositions for possible content. To increase the motivation, no requirements were made concerning the content. Students were encouraged to work on proper projects. For those without creative ideas a list of propositions for possible content was given.

An example of an exercise for digital image processing with Photoshop can be found in the appendix.

3. Deliverance and Declaration of a document

Every *author* was supposed to upload his *document* via a web-interface before the date of delivery and therefore declare his participation for the given exercise. The delivery of a document was binding. No corrections (like delivering a second document overwriting the first) were possible once a document has been uploaded even if the date of delivery was not yet reached. The learning goal was the preparation for similar situations in professional projects or job applications. Once sent further corrections are impossible.

A note could be added to a document, explaining particular design decisions and helping the reviewers in creating their reports. A student capable of planning, describing and justifying his workflow could consequently influence the evaluation of his document.

Every author who declared his document properly was assigned deep links to five other documents for reviewing and evaluation. Every delivered and declared document was therefore reviewed by five authors considered to be experts since they worked on the exercise and knew the particular problems. Five reviews should assure that the differences between the experts would statistically compensate at the end of the semester.

Since the documents were distributed among the authors in a way, that every document was to be reviewed by exactly five experts, delivering a document after the date of delivery was not possible for this would have meant to reshuffle all review tasks.

4. Evaluation of documents

A document could be evaluated with up to three points:

- 0 points: No reference to the exercise visible or document not readable
- 1 point: basic work
- 2 points: exercise fulfilled and the formal requirements were respected
- 3 points: requirements surpassed

Additionally a short review was to be written to justify the evaluation using the terminology of media design presented in the practical courses.

Being the tutor I hold the option to interfere into the process of reviewing if the difference between several reviews was too large. Normally the points of each evaluation were summed up to the overall score of the document.

The deliverance of all five reviews were crucial to validate one's own points. A reviewer handing over less than the five demanded reviews would have received no points at all for his document. To prevent disadvantaging an author in case of missing reviews and points for his document, the overall score was filled up with the rounded average of the points of the existing reviews.

Example: A document *mydoc* has been reviewed by four experts and evaluated with 2, 2, 3 and 3 points. The last review has not been delivered in time. The author of *mydoc* has to review five documents himself. As soon as this has been done he gets an overall score of $2+2+3+3+\text{round}(2.5) = 13$ points. If he misses to deliver one or several reviews, he gets no points at all. This was enough incentive for the experts to deliver the reviews in time. Exceptions due to illness or personal accidents

could be negotiated with me and the points could be nevertheless validated afterwards.

The life-cycle of an exercise consists of five distinct parts: (1) *presenting the exercise*, (2) *delivering the documents*, (3) *distribute the review-tasks*, (4) *delivering the reviews*, (5) *calculating the scores for the documents*. Counting a week for each exercise (we actually accorded two weeks) the time-structure would be as followed:

- **Week 1**
Tutor presenting exercise 1
- **Week 2**
Students delivering document 1
Tutor distributing the review tasks for document 1
Tutor presenting exercise 2
- **Week 3**
Students delivering the reviews for document 1
Tutor calculating the scores for document 1
Students delivering document 2
Tutor distributing the review tasks for document 2
Tutor presenting exercise 3
- ...

In a given week usually three exercises are in the pipeline which can be easily managed with a detailed documentation assuring that the students know exactly what to do.

Additionally a ranking list has been offered. With this dynamically generated list every student could anonymously compare his score to those of others.

5. Technical Framework

The review-system was programmed in PHP in connection with a MySQL-database. The most sensitive script is the management of the reviews and scores. It was mandatory to prevent a reviewer to evaluate his own document or that of a friend. To discourage the students from testing the system, every request that could be interpreted as fraud was saved with the corresponding IP-address. Since the students did not know how these datas were treated, systematic intrusion attempts could be prevented. Of course, every request has been properly filtered to protect the system against SQL-injections and similar attacks. During the practical course no successful attack nor a result of such an incident could be documented.

Besides the scripts forming the interface to the students, the central script was a table showing the delivered documents, the review tasks with links to the delivered

reviews and the results. Problems like missing documents, missing reviews or large differences between evaluations could be easily illustrated.

After a second application during the semester 2005/2006 the scripts will be documented in detail and published as open source.

6. Experiences

The students got quickly be used to the review-system. After the deliverance of each review the number of required reviews and the validation status of the proper points were displayed, helping to deliver the reviews in time. As already mentioned, no problems showed up due to missing reviews.

To evaluate the system, a questionnaire has been handed to the students at the end of the course interrogating crucial parameters. It aimed at establishing a representative opinion on which would depend the decision if the system is to be used in future courses. The results are very encouraging, the great majority of the comments were neutral or positive. One of the most important reasons for this result was the transparent description of the scale of evaluation. The impression of getting comparable points by different experts is based on this description. As a consequence the tutor must try to describe the learning goals as objective as possible.

The review-system has one serious flaw: The reviews are themselves not reviewed. Therefore it is possible for the students to write effortless content-free rapports. It actually happened: *«Very nice, 3 points», «Nothing to complain! Top! 3 points»* or *«Nothing more to say I think ... 3 points»*. An explanation for the three points – i.e. maximum points for documents that most probably hasn't even been opened (in one of the cases the document was technically corrupted) – is solidarity amongst students. Maybe the demand to evaluate other students was considered as somehow unethical by these authors. Content-free reviews like those mentioned are easily discovered (`SELECT * FROM reviews WHERE LENGTH(review)<50 AND points=3`) but even a brief inspection of all reviews is quickly done with appropriate scripts. After threatening a loss of points for authors writing content-free reviews the problem disappeared.

Two reproaches were commonly repeated in the questionnaire:

First, the choice between 0-3 points was judged too limited. Mostly, a range between 0 to 5 points was asked for. This scale is known from german schools with points from 1 to 6. This demand will be respected the next time the review-system will be put into action.

Second, the peer-reviews were considered unjust. This correlated strongly with a gap between auto-perception and evaluation. Some students claimed their documents to be underestimated by incompetent reviewers.

Dealing with this reproach is more difficult. It is based on the impression, that students are not qualified enough to evaluate their fellow students. Although review tasks were sent exclusively to students who worked on the corresponding exercise, who delivered a document and who were consequently familiar with the problems and issues, what about authors who worked only basically on a document? How can they judge an elaborate work? Fortunately, few of these problems occurred, which was mostly owed to the character of the exercises – media design with standard software can be judged even without knowing too many details. Most of the students considered their documents fairly judged, some felt underestimated but some admitted to have earned too good scores.

7. Applications

Although the review-system has been designed for academic teaching, other applications are possible. For further explications it reveals helpful to distinguish the roles concerned in a practical course: There are the roles of the students as well as the didactical roles, decreasing in the order of their didactical qualification: teacher, exercise designer, tutor and corrector. Following a modern didactical approach *students* are people seeking help while learning. *Teachers* for their part help students to learn. The *exercise designer's* qualification is the didactical transformation of subject matter into exercises. The role of the *tutor* is to present the exercises to the students. The *corrector* reviews the students' works following the instructions of the exercise-designer. In school all didactical roles are taken by the teacher. In university courses with many participating students, the roles are often distributed between the professor (teacher), assistant (exercise designer), graduate or undergraduate but experienced students (tutor and corrector). The latter are responsible for groups of 10-20 students. This is acceptable for most of the students having somebody a bit more experienced to ask for help. The limiting factor for the size of a group is the effort that can be taken by a single corrector. From a didactical point of view the work of a corrector consists of giving the students a helpful feedback. But more often than not his work is constraint to marking mistakes guided by the solution model of the exercise designer and counting the score without giving any further remarks. The role of the corrector hence requires the least didactical qualification and is therefore the least paid job yet consuming the largest amount of time.

The advantage of the review-system lies in assigning the role of the corrector to the students. The most time consuming work, at the same time requiring the least didactical skills, is taken from the scarce and therefore expensive didactical personnel. But it would be too short-sighted to note only financial benefits. To economize a resource opens new possibilities in its distribution. The size of a learning group may increase if personal feedback of a tutor is less important. In the course documented

in this article one exercise designer and tutor managed a learning group of 90 students in a reasonable amount of time.

The system may be used in adult training where it is easy to find and finance a teacher offering lessons and designing exercises but who is unwilling to correct the students work. The review-system is the tool of choice if (i) the score attained in the exercises is not necessary to obtain a grade and (ii) feedback is important but can't be offered by a tutor.

Completely virtual training centers are possible, e.g. for software, often necessitating intensive training. Online-communities can form around a given subject getting their training exercises from experienced users but managing the evaluation themselves. Such online-communities may stay completely anonymous in which case the registration name is replaced by a pseudonym and the registration number is assigned by the system during inscription.

8. Concluding Remarks

The review-system is a tool for Computer Aided Instruction (CAI) but not adapted to all kind of learning situations. It seems well designed for exercises demanding not too much qualifications to correct. It was widely accepted among the students who were surely aware of the fact that they would get far less feedback for their work without it. But it requires cooperation from the students often corresponding to a certain level of education. It equally requires a group size larger than 30 participants to assure the anonymity of the reviews.

9. Appendix

Example of an exercise

Underlined text was originally a link to a file or a website offering in-depth information. The propositions for content promoted projects trying to develop cultural activities on the campus but needing support in graphic design. During the course some very good works were produced by the students.

Exercise 2: Image processing with Photoshop

Date of deliverance: 22.01.2004

File-forma: Photoshop document (PSD) or ZIP-archive.

Document-name: *[student-number].psd* e. g. *7533428.psd*

Formal requirements: Design a Photoshop document by combining several bitmap images (e.g. surfin' Duck.psd, 1.4 MB). Don't let the document become too large. It should include

- At least 3 layers
- Selected elements with the selections saved as alpha channels (e.g. duck.psd)
- Layer masks
- Optimized levels and curves.
- At least one of the following techniques:
 - Knockout
 - Clone stamp tool
 - Duotone
 - You may moderately use layer styles

Possible content

- Design a poster for one of the following projects:
 - Hello World!
 - Spätlese
 - Hörspielwerkstatt
- Design a website, maybe for one of the noted projects. The planning of a website with Photoshop is standard in the workflow of professional web-design (e.g. piccolloSite.psd, 2.2 MB).

- Design a comprehensive photo montage. Examples may be found in the exhibition “Lying images” at the German Historical Museum.
- Get inspired by the tutorials on the FTP-Server in the folder *digimedi.imageprocessing.tutorials*.

Further sources:

- Image-Galleries are on the website of the practical course
- Google offers a sophisticated search function for images
- High-quality tutorials can be found on the website of Carl Volk
- Another good source for tutorials is PlanetPhotoshop
- Of course, Adobe offers tutorials as well

Happy designing!