STUDENTS' PERCEPTION OF E-GRADING: INFORMATICS COURSE CASE STUDY

Arkadiusz Naruk[1]

1. Wydział Elektrotechniki i Automatyki
    Politechnika Gdańska
    tel.: 58 347 13-32   e-mail: anaruk@ely.pg.gda.pl

Abstract: Using computers for grading can simplify work for teachers. However it is a sensitive part of their work and must be performed carefully. More information is required on students' reception and opinions about e-grading. This paper is a case study of e-grading used in Informatics course and discusses results obtained from surveying the students after being graded by a teacher and a computer.

Keywords: computer grading, educational technology, electronic learning

1. INTRODUCTION

Growing and increasingly more popular e-learning has had to lead to emergence of more e-based aspects of teaching. E-grading or computer grading is among the most important of those, because it regards crucial part of education - knowledge or skill verification. One cannot say if learning was successful or not without proper verification. It is a risky process as tests, quizzes and other assignments have to be adequate to the subject, but moreover, grading is a very sensitive part of learning. It finalizes the process of learning, which often affects future effects like improvement or decrease in student's performance [1]. Fair grade which met student's expectations can make stronger positive connection to the material and help student remember the subject better [2].

There is another benefit that comes with computer-given grades - for obvious reasons it enforces an assignment to be in a digital form. This opens ways for numerous processing methods of material uploaded by students. For example, computer allows for cross-checking all assignments for hints of plagiarism, which is beyond human practical capabilities. Computer can compare all uploads with each other and provide information on similarity level or similarity percentage. Student homework assignments or tests that are written is some formalized language (such as programming or mathematical language) can also be broadly analyzed and verified - e.g. program source codes can be compiled and then run. Moreover, essays can be checked from more of technical point of view, such as counting number of words used which could give an idea about vocabulary richness in just one click. Such tasks would be otherwise practically undoable by a human alone.

The above combined give e-grading a reason to develop and make it desirable tool for teachers nowadays.

However, computer grading is still being studied and its impact evaluated. This paper presents results of a study which purpose was to gather students response to two ways of grading - done by computer and teacher, and perform comparative analysis of them in light of possible flaws and caveats for future use. Paper shows there is still research potential in the field.

2. STATE OF THE ART

There has been a considerable number of efforts to use computer grading more often and more intensively, because of tempting pros which it brings into e-learning. Many aspects have already been studied or at least mentioned in the studies. First of all, teacher is always susceptible to mood changes, student's affection and such, which determine the way he hands grades out, while computer grades are always given according to the algorithm and will never change on their own. That consistency gives chance for fairness in the classroom which is very important [3] as shown in [4] and [5].

Secondly, computer grading relieves teacher of this tedious and time-consuming task [6], so he might devote more time e.g. to prepare good quality material for students rather than burn time checking their work. However, it may seem to be at the expense of detailed and individualized error corrections [6], but research (e.g. [7], [8], [9]) shows that '*well-designed computer feedback can be more effective than manually-graded homework assignments in producing significant differences in learning*' [6].

3. CONTRIBUTION

Research shows (as in e.g. [1]), grading can have both negative and positive effects on students. Meddling with grading methods is a dangerous field as it is a sensitive area of teacher's work. Therefore more information is required how students perceive computer-given grades and 'hybrid' systems which consist of combined computer and teacher marks at once.

The main scope of this paper is to grasp major students feedback on being graded by the computer and their opinions on the topic. Specifically, it aims to provide information on:

*   How do students see fairness of computer grades compared to teacher's grades?

- What would change if computer grade they got was given by a teacher?
- How do students think they should be graded?
- Which of grading do students prefer?

The main contribution of this paper is basic comparative analysis of various aspects of two grading methods and conclusions about students perception of them. Possible new areas of future research are included in the end.

4. PERFORMED STUDY

Studied subject was Informatics course for undergraduates of third semester of Automation and Robotics. During the course students were to do a semestral project which was to write a computer game. Students were given specific project requirements which would be taken into account during grading such as number of animated objects, possibility to pause the game, save and load options, difficulty levels. Besides the project itself, they were to hand in progress reports (short descriptions of work done in a given week) and current version of game source code. Course was hosted on MOODLE platform and students uploaded their work there. There were specific time limits set for each upload. Students were to hand in their report in text format and an archive with a source code.

Four things were graded:
- reports,
- regularity,
- game functionality,
- source code quality.

The first couple was graded by a computer, the second – by a human teacher.

In case of reports, it was enough to upload them. Their quality was not controlled. Only percentage of reports uploaded was taken into account and directly converted to a grade received by a student.

Regularity measurement was more complex. There was a Linux shell script and C program processing all of the student uploads. Source code from n-th week was compared with n-1-th week in terms of number of changed lines. To describe student's progress, four-level scale was used to describe the level of changes with possible options: *almost unchanged*, *minor change*, *fair change*, *major change*. Assigned ranges were based on experience from previous years of the course. Based on these descriptive grades, one numerical grade was counted and given as 'regularity grade'.

Game functionality means a number of project requirements met, as described above e.g. number of animated objects or possibility to save the game state. Grading was done by the teacher.

Source code quality was the most subjective component, graded also by the teacher. It included code clarity, program architecture (class or function organization), variable naming consistency, comments and other things which can be seen as 'good programming practice'.

Each of the grades was in range of 0÷100%. On pre-last meeting students learned their computer-given grades. During the class, the teacher checked their games and graded them.

On the last meeting, students were surveyed for how they see those grades. The survey contained several single and multiple choice questions. Total of two project groups were surveyed, total sample size was 33 people. Results of the survey are discussed in chapter 5. Graphs sometimes omit least chosen answer.

5. RESULTS

First analyzed question was about perception of justice in dependence of satisfaction from the grade given by both computer (Fig. 1) and teacher (Fig. 2).

First graph shows students' votes distribution, whom computer graded lower than would satisfy them. Difficulty in justice evaluation of such grade can be seen in similar votes distribution between '*just*' (22%) and '*unjust*' (28%), but also dominance of '*do not know*' answer (50%). Second histogram shows satisfied students. It can be observed that at least some of them identify grade's justice with personal satisfaction – 75% of them thinks that the grade is unjust, whereas only 25% remains undecided. Only one person got more than would be satisfactory, which can be assumed as an error margin and neglected in the analysis. It can however serve as an information that system was not too mild, if students got at best what would satisfy them and not more.
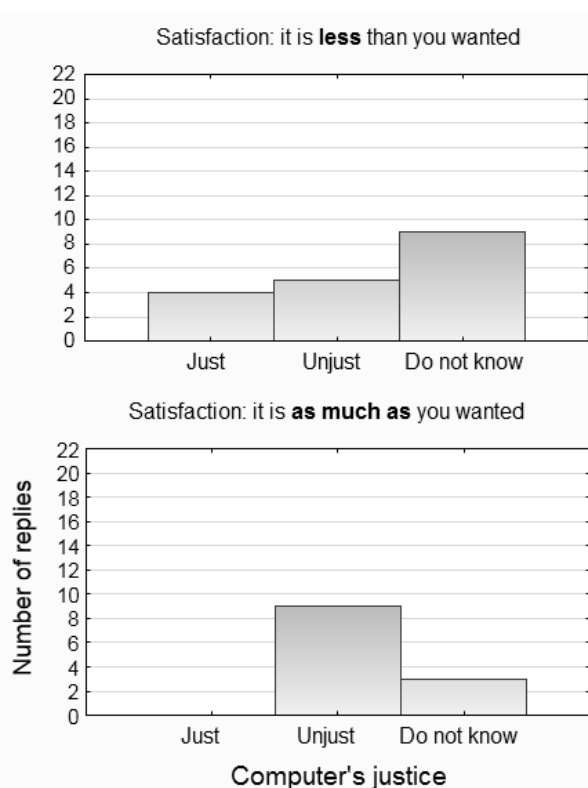

Fig. 1. Histograms: perception of computer's justice in dependence of student's satisfaction level

In case of teacher's grade, perception of justice was significantly different than in case of computer grade. Number of students who got lower grades than would satisfy them was considerably smaller and despite this, none of them saw their grade as unjust. Majority of students got what expected (70%), whereas 87% of them saw the grade as just. Results lead to conclusion that despite the grade, student assumes teacher's grade as fair, however this test would have to be repeated on bigger sample. It is also surprising that nobody chose '*unjust*' option – in worst case students marked '*I do not know*', despite the fact that survey was anonymous and teacher did not control or see chosen options.

Satisfaction: it is **less** than you wanted

Satisfaction: it is **as much as** you wanted
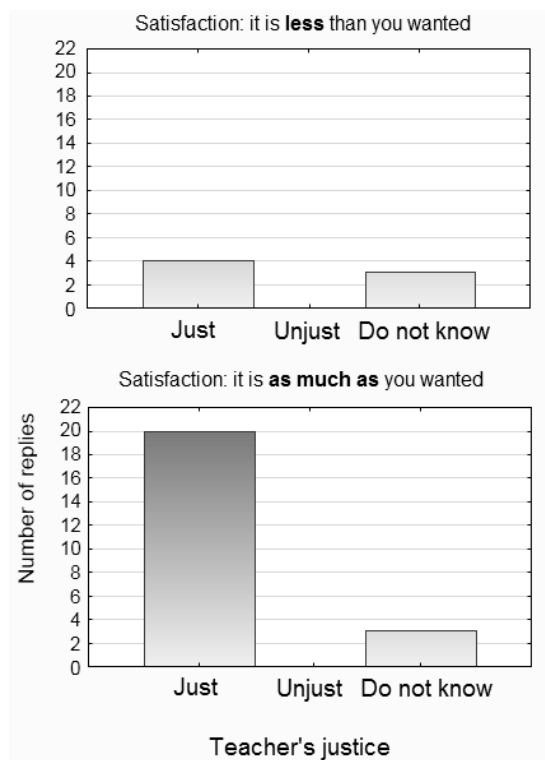
Teacher's justice

Fig. 2. Perception of teacher's justice in dependence of student's satisfaction level

Another question which gave significant results was *'What would you do if this grade was given by a teacher?'*. From 30 students who answered this question 66,6% would try negotiating. Fig. 3. presents histograms which show dependence of the answer from satisfactory grade. One can see that unhappy students would try negotiating the grade if it was given by a teacher, not by a computer. In case of satisfied students, almost half of them would also try negotiating. Therefore it can be concluded that teacher's grading can be adjusted to personal needs of students, but it is definitely vulnerable to exploits.

Another goal of this survey was to find out how students think they should be graded. Emphasis is on the word 'should', which in this context means '*done in an appropriate way*'. It is interesting that just 34% responded '*by human teacher*', while over 54% say that grading should be hybrid, mixing teacher's and computer's mark. Only 1 person claims that grading should be handled just by a computer, which can be assumed as marginal value and neglected. Histogram showing relation between student's grade satisfaction and answer to discussed question is shown in the Fig. 4.

In the Fig. 4. (a) and (b) histograms are based on students' computer's grade satisfaction level, while (c) and (d) are for teacher's grades. It is notable that on 3 out of 4 graphs (b, c, d), hybrid way of grading is dominant answer and in the remaining case (a) it is almost as popular as the winning answer. It is important to notice that case (a) naturally encourages students to choose answer '*teacher*', because students in this group got less than they wanted. Despite this, answers are split almost evenly.

This is interesting, because despite all the doubts and opinions about computer grading, students think it *should* still be part of the grading system. They however do not prefer this when given an option to choose. Almost 70% of students answer '*teacher*' to the question '*Which of grading would you prefer?*', while negligible 3% say

'computer'. 27% of students claim that there is no difference. Connection of this question and computer grade expectations is shown in the fig. 5.

Satisfaction: it is **less** than you wanted

Satisfaction: it is **as much as** you wanted
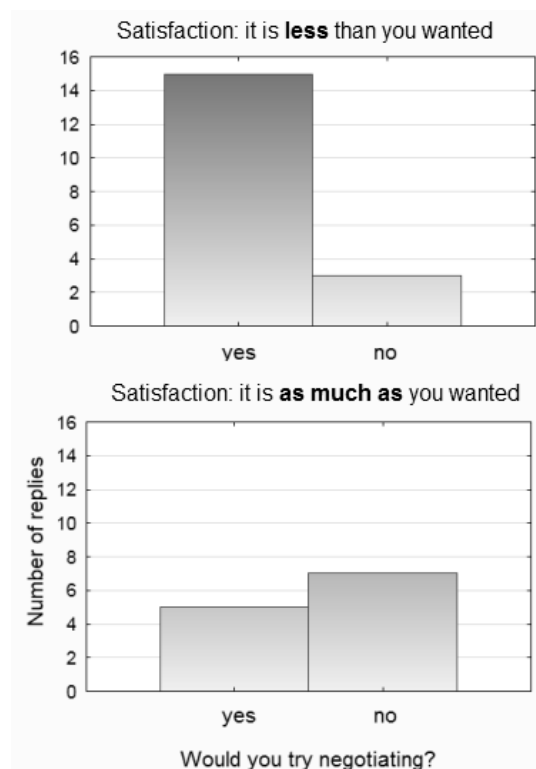
Would you try negotiating?

Fig. 3. Results from a question *'What would you do if this grade was given by a teacher*: I would try negotiating the grade' in dependence of student's satisfaction level
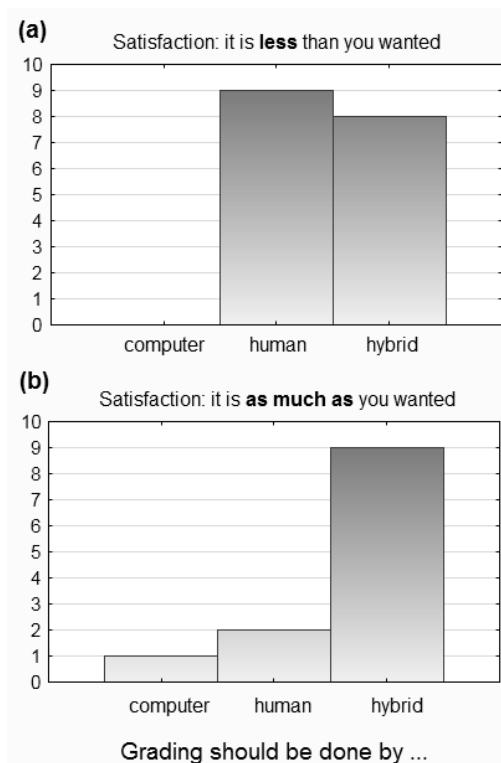
(a) Satisfaction: it is **less** than you wanted

(b) Satisfaction: it is **as much as** you wanted

Grading should be done by ...

Fig. 4. (a), (b): histograms of a question *'Grading should be done by ...'* in dependence of student's satisfaction level

(c)

Satisfaction: it is **less** than you wanted

(d)

Satisfaction: it is **as much as** you wanted

Grading should be done by ...
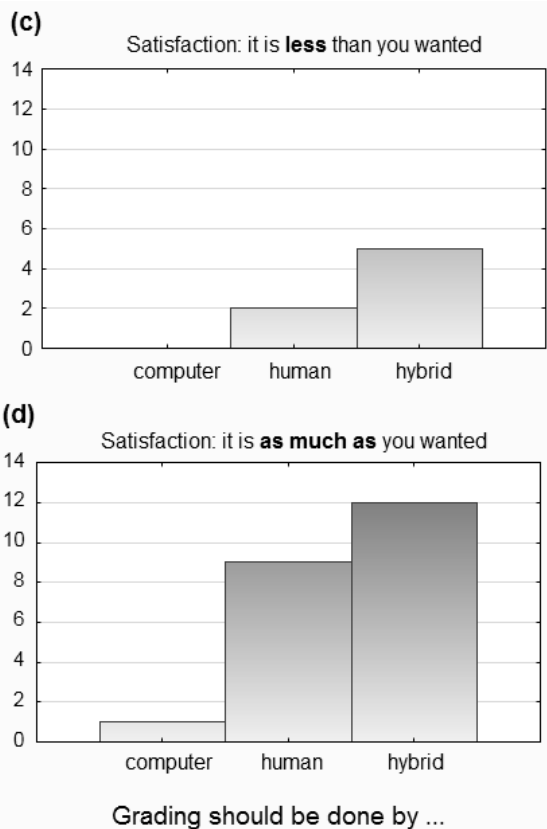
Fig. 4. (c), (d): histograms of a question *'Grading should be done by ...'* in dependence of student's satisfaction level

Satisfaction: it is **less** than you wanted

Satisfaction: it is **as much as** you wanted
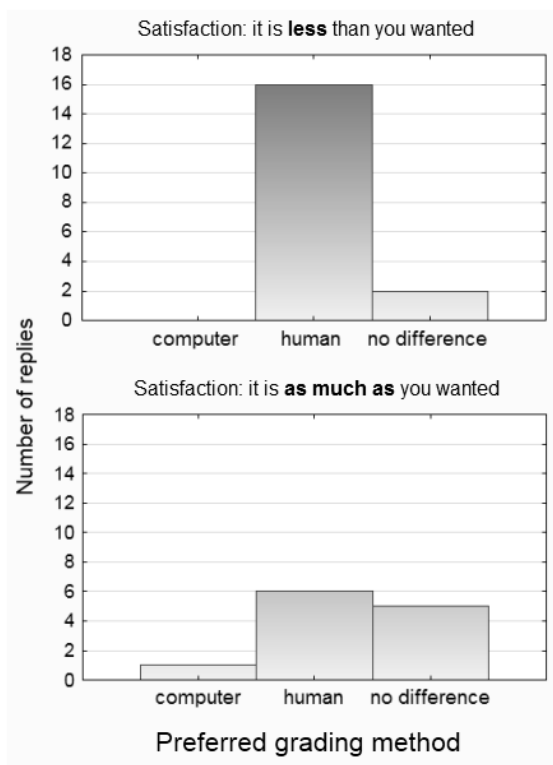
Preferred grading method

Fig. 5. Histograms of a preferred grading method in dependence of student's grade satisfaction level

Results presented in the Fig. 5. are logical - if computer grade is lower than expected, students automatically prefer to be graded by a teacher

6. CONCLUSIONS AND FUTURE WORK

This study shows that measuring students perception of grading is a difficult task. Students are not unequivocal in their interpretation of questions. Number of blank fields in the survey shows that students have hard time deciding what they actually think and feel. It was however possible that judging computer's fairness is hard, because teacher is naturally closer to students and they automatically prefer him over computer, even if they cannot explain or justify this.

The study also shows teacher's grading vulnerabilities. The most important is that students would try negotiating the grade if it was given by a human teacher instead of a computer. Conclusion of this could be a hint for teachers to be aware of this and try to remain resistant to such attempts.

One unexpected result of the survey was divergence between student's opinion and personal preference regarding computer grading. Students mostly agree on the fact that grading should be done in hybrid way (by both computer and human teacher), but when given an option, they prefer just teacher's grading. What advantages or needs did they actually see behind it remains unclear and can be studied in future work.

Students often seem to connect personal satisfaction with perception of justice, but this is beyond scope of this paper and remains a task for future research. However, based on this survey, students prefer being graded by a human teacher and think it is a fairer way.

This study also reveals a few possible areas for future exploration. Firstly, it is unclear why students think that grading should be done as a mix of computer and human work. Source of this view might be a valuable thing to be checked. Another aspect of this is why exactly students prefer being graded by a human teacher? Does this depend on a subject? Is it so that it was their only time and the system was somehow flawed, but would be preferred otherwise?

Secondly, one can measure relationship of these results with time. As e-learning expands and more electronic tools are used, opinions of students who are more accustomed with those may get more positive in time.

Thirdly, this paper touches topic of justice just slightly, however it may be worthy to evaluate the matter deeper. Plain students' justice level can be evaluated on its own, but also in comparison to students' satisfaction levels. Putting justice versus satisfaction might also be capable of producing noteworthy results.

Last but not least, suggested ways are in regard of students' perception of computer grading. But how to objectively measure computer grading on its own? Is there a common denominator between two computer grading systems that would allow for their comparison? These questions open possible new areas for further research.

1. Lipnevich A. A., Smith J. K.: *Response to Assessment Feedback: The Effects of Grades, Praise, and Source of Information*. ETS, June 2008. Available 23.01.2014 at http://www.ets.org/Media/Research/pdf/RR-08-30.pdf.
2. Surgenor P.: *Teaching toolkit. Effect of Assessment on Learning*. University College Dublin, January 2010. Available 20.01.2014 at http://www.ucd.ie/t4cms/ucdtlt0031.pdf.
3. Horan S. M., Chory R. M., Goodboy A. K.: *Understanding Students' Classroom Justice Experiences*

and Responses, Communication Education, 59:4, 453-474, 2010.

4. Moore, M. L., Moore, R. S., & McDonald, R.: Student characteristics and expectations of university classes: A free elicitation approach. College Student Journal 2008, 42, 82-89.

5. Walsh, D. J., & Maffei, M. J.: Never in a class by themselves: An examination of behaviors affecting the studentprofessor relationship. Journal on Excellence in College Teaching 1994, 5, 23-49.

6. Dabrowski R., LeLoup J. W., MacDonald L.: Effectiveness of computer-graded vs. instructor-graded homework assignments in an elementary spanish course: a comparative study at two undergraduate institutions. IALLT Journal of Language Learning Technologies, Vol. 43 (1) 2013, p. 78-100.

7. Nagata, N.: Computer vs. workbook instruction in second language acquisition. CALICO Journal 1996, 14(1), 53-75.

8. Hubbard, P., & Siskin, C. B.: *Another look at tutorial CALL*. ReCall 2004, 16(2), 448-461.

9. Godwin-Jones, R.: *Emerging technologies – focusing on form: Tools and strategies*. Language Learning & Technology 2009, 13(1), 5-12.

ODBIÓR E-OCENIANIA PRZEZ STUDENTÓW: STUDIUM PRZYPADKU NA INFORMATYCE

Wykorzystanie komputerów do oceniania może uprościć pracę nauczycielom. Jednakże jest to wrażliwa część ich pracy i musi być wykonywana z ostrożnością. Potrzeba informacji na temat odbioru i opinii studentów na temat e-oceniania. Niniejszy artykuł stanowi studium przypadku wykorzystania e-oceniania na kursie Informatyki i omawia wyniki ankiet studenckich, którzy ocenieni zostali przez nauczyciela oraz przez komputer.

Słowa kluczowe: komputerowe ocenianie, technologie edukacyjne, elektroniczne nauczanie