Procedia Computer Science

Volume 80, 2016, Pages 1812–1821

ICCS 2016. The International Conference on Computational Science

The Impact of Learning Activities on the Final Grade in Engineering Education

Raul Ramirez-Velarde[1], Nia Alexandrov[2], Miguel Sanhueza-Olave[3], Raul Perez-Cazares[1]

[1]*Tecnologico de Monterrey, Monterrey, Mexico*

[2]*Barcelona Supercomputing Center, Barcelona, Spain.*

[3]*Universidad Tecnologica Metropolitana del Estado de Chile*

rramirez@itesm.mx, nia.alexandrov@bsc.es, msanhueza@utem.cl, raul.perez@itesm.mx

Abstract

This paper investigates the impact of the learning activities on the final grades in students in engineering education. A principal component analysis is carried out on the undergraduate "Stochastic Models" course. We have determine that the first principal component has a positive correlation with the score of the final written cumulative exam. This could possibly mean that the final exam could be eliminated from the engineering curricula, but the variability is significant as measured by the correlation R statistic. Based on information gathered on a much larger sample, we found that the variability has increased, indicating changes in the course and students' emphasis in learning activities. Therefore it can be concluded that the evidence presented does not justify eliminating written cumulative final exams.

Keywords: Principal Component Analysis, Course Design, Learning Activities, Evaluation, Learning Analytics

1 Introduction

In education, the necessity of homework, exams and other learning activities has been questioned by teachers, parents and students for a very long time. Some people believe homework is unnecessary while others believe they are indispensable. The same can be said about written cumulative exams. While few people question the necessity and validity of partial exams, many question the need and/or the value of final exams. Recently, in (Ramirez-Velarde, Sanhueza-Olave, Alexandrov, & De Marcos-Ortega, 2015) we were able to design a statistical experiment that allowed us to determine if learning activities were correlated with the grade obtained in a final cumulative exam. We concluded that if a course was correctly designed, students that did their learning activities achieved higher score on the final grade. Nevertheless, the side effect of such study is that it may show that final exams are unnecessary. That is to say that if school work could be correlated to the score on the cumulative final exam, then we could eliminate the final exam altogether. Could this be so?

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2016

doi:10.1016/j.procs.2016.05.465

For example, the study (Maltese, Tai, & Fan, 2012) shows very weak correlation between homework and final exam scores at K-12 level, even in science and mathematics. Clearly this shows that under correctly designed exams, some learning activities could be eliminated. On the other hand, there is the issue of eliminating the cumulative exams, not the learning activities. For example, recently Harvard University announced that cumulative final exams were no longer mandatory (Harvard, 2010), and in (Wiggins G. , 2014) it is argued, that under certain circumstances, for example when final exams favour content mastery more than understanding, final exams should be eliminated.

To continue the discussion, let us define learning and let us establish guidelines for learning activities and cumulative final exams. We do that in section 2. In section 3, we revisit previous results and discuss them. In section 4 we present new statistical analysis and we argue in favour of the need for final exams and in section 5 we draw our conclusions.

2 Definition of Learning and Final Exams Grades

In (Wiggins G. , 2014), it is argued that sometimes, course design is expressed as a dichotomy between content mastery and understanding. Therefore, some evaluation instruments such exams, would be oriented to either. If the course orientation is content mastery, then final exams can of course be eliminated, as it is argued that content mastery is only a tool of understanding. In (Wiggins G. , 2014) (Wiggins G. P., 1998), is stated that students that:

possess content mastery are capable of	have achieved understanding are capable of
- Recall concepts - Repeat practiced activities - Perform as practiced/be competent - Plug in practiced knowledge - Recognize/identify what they have learned	- Justify a claim - Connect discrete facts on their own - Apply their learning in new contexts - Adapt to new circumstances - Criticize arguments made by others - Explain how and why

And of course learning evaluation instruments would be designed following those aimed abilities. It could be entirely the case that if we can eliminate final exams from mastery oriented courses is because poor course design, not because final exams per se are unnecessary.

2.1 Definition of Learning

In our definition of learning, perception and memorization are only a part of the whole, and not even the first part. We define learning as a progression of increasing knowledge complexity that follows the path: Conceptual and Contextual Knowledge->Procedural and Problem Solving Knowledge->Cognitive complexity knowledge. Table 1 shows the conceptual levels, and Fig. 1 shows the learning cycle (Alexandrov & Ramirez-Velarde, 2013).

Notice that in our definition, learning is a cycle, and therefore, it has no beginning. In traditional learning, perception (concrete experience) and memorization (construction) would be first. But as a cycle, learning can start with action, like in interactive museums, with construction, as in the Socratic Models or even with abstraction by stablishing hypothesis first using creativity such as in (Ramirez-Velarde, Perez-Cazares, Alexandrov, & Garcia-Rueda, 2014).

Specifically, the transition from construction to abstraction requires transformations of knowledge that must take place in order to achieve deep understanding and competence. This should be considered the core of the learning process (Alexandrov & Ramirez-Velarde, 2013):

1) **From past to future**. Information given to students is by nature the past. Students must be able to make plans and create strategies: that is to project past learning into the future. This is effectively achieved when we transition from reflective observation to abstract hypothesis. Observations are in the past, therefore we reflect about those observations. Plans and hypothesis intend to predict the future.
2) **From inside to outside**. Human beings receive knowledge through their senses. It must then be effectively stored. After such process, new knowledge is created, transforming students from knowledge receivers to knowledge producers. Again, this is achieved in the transition from reflective observation to abstract hypothesis. We store inside our minds, in our memory, knowledge that we took from the outside. When we make plans or create hypothesis, we project what we know and create new knowledge about the real world, the world outside our mind. This the transformation of knowledge that gives the strategy its name.
3) **From learning to teaching**. This is a power transformation in which initially students are dependent on outside authority to inform them. Eventually, students take control of their learning taking decisions of how, where and why. Teachers become tutors and even mentors, through a carefully constructed scaffolding lattice of slowly retracting learning support, until student become experts on the knowledge area.

Knowledge Level	Learning Activities
Conceptual and Contextual	Students read and see explanations Students create and discuss the solution of questionnaires based on memorization Students design questions to engage in discussions
Procedural and Problem Solving	Students practice with examples, worked exercises and complex problems Students design exercises and problems Students learn how to create questions working together and how to discuss the solutions
Cognitive Complexity	Students use their creativity to adapt solutions to new problems Students use their creativity to solve new problems with no known solution Students not just answer questions, they make them. The questions have to be unique and with appropriate complexity or difficulty

Table 2. Taxonomy of items to be included in questionnaires

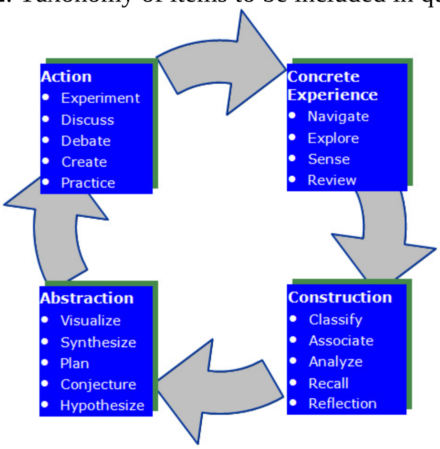

Figure 1. Integrated Learning Processes Educational Model

2.2 Evaluating Learning

According to Kirkpatrick there are four levels in an evaluation system (Kirkpatrick, 1996):

1 **Reaction of learner**. What they thought and felt about the education and training
2 **Learning**. The resulting increase in knowledge or ability
3 **Behaviour**. Implementation of increased capability in a particular activity. Also called Transfer
4 **Results**. The effects on the business or environment resulting from the learner's performance

In (Alexandrov & Ramirez-Velarde, 2013) we call level 2 cognition and levels 1 and 3 metacognition. Level 4 is related to competencies. Each of the levels should be measured for full meaningful evaluation. In this article we will concentrate on levels 2 and 3, that is, cognition and metacognition.

How should evaluation of knowledge and behaviours be carried out? Evaluation is such an important aspect of our learning systems that we should try to design it with the least amount of bias and error. Svinicki suggests that we should try to achieve the 4 R's (Svinicki, 1999) (Pointek, 2008)

1 **Relevant**. Or **Validity**. Providing useful information about the concepts they were designed to test. Any activity used to evaluate a student's learning must be an accurate reflection of the skill or concept which is being tested
2 **Reliable**. Allowing consistent measurement and discriminating between different levels of performance. That means that sufficient information about required performance should be given, instructions should be clearly communicated and the evaluation criteria should be static. It means that students will equal skill will get equal grade.
3 **Recognizable**. Instruction has prepared students for the assessment. Students should be aware of how they will be evaluated and their class activities should prepared them for those evaluations.
4 **Realistic**. Concerning time and effort required to complete the assignment. That is to say, that the amount of information obtained is balanced by the amount of work required.

So, how should learning evaluation be carried out? As pointed by Pointek (Pointek, 2008) there are many instruments for evaluation. Nevertheless, two characteristics stand out:

- Valid, reliable and recognisable evaluations should include the verbs: explore, review, classify, associate, analyse, recall, reflect, synthesize, hypothesize, discuss, debate, practice, solve, create, etc.
- As pointed by (Myers & Myers, 2006), evaluation, specially written cumulative exams, should be continuous and as frequent as possible, since this improves performance, improves student satisfaction and reduces anxiety. Students that are evaluated frequently obtain better scores in cumulative written finale exams than students that are evaluated only twice in a school period.

In this paper, we will concentrate on cumulative written final exams. Our aim is to determine the validity, reliability and recognisability, of those in university level education. In order to determine validity, we will carry out a statistical analysis based on principal component analysis (PCA) similar to the one carried out in (Ramirez-Velarde, Sanhueza-Olave, Alexandrov, & De Marcos-Ortega, 2015) to see if the work carried out by students to prepare for the evaluation allows them to get satisfactory grades. If students carry out activities but that is not reflected in a satisfactory score on an exam, then course activities are not relevant or the exam is not relevant. We will also evaluate if school work factor does discriminate between students that carry out those activities and those who don't in the score of the final exam. Also, if students activities do prepare them for their evaluation, then the evaluation is recognizable, as students understand how they will be evaluated.

3 Multifactor Analysis of School Work and Scores on Final Grades

In (Ramirez-Velarde, Sanhueza-Olave, Alexandrov, & De Marcos-Ortega, 2015), we analyzed several courses by carrying out a principal component analysis (PCA) trying to establish if learning activities were correlated with the score of a final cumulative written exam. We found that the square cosines measure associated the first principal component (PC1) with most of the learning activities, as shown in table 1 for the undergraduate course Stochastic Models taught from January to May 2015. This means that there is only one path to getting a good score on the final test. If activities are associated to a different principal component, then it could mean that a student can follow different paths to a good grade. For example, a student can have low scores in some activities associated to certain principal component but if she carries out a set associated to another principal component, she might still achieve a good score on the final exam.

In figure 2 we see a positive relation between the first PC and the final exam score with a linear fit with R statistic of 0.634, meaning that the linear fit represents 61% of the data. We concluded that if a course is correctly designed and evaluated, then there must be only one path to a high score on the final exam and that path is to carry out all learning activities. Nevertheless, there are some activities that are so important in the course that can be separated by their own PC. This is again illustrated with the Stochastic Models course in figures 3 and 4. In this course, Homework about M/M/1 queues and, about probability and random variables are actually considered by students the main learning activities of the course. The attention and effort given by students to these activities separates them into the second and third PC. Nevertheless, although there is correlation between the mark obtained in these activities and the final exam, the relationship is weak. The explanation of this requires further study.

Activity	F1	F2	F3
Par 1	**0.453**	0.015	0.166
Par 2	0.361	0.007	0.006
Par 3	**0.521**	0.045	0.084
Sim 1 y 2	**0.610**	0.140	0.065
Sim 3 y 4	**0.478**	0.306	0.004
Prob	0.095	0.067	**0.632**
Quiz	**0.342**	0.115	0.081
M/M/1/∞	0.301	**0.464**	0.184
M/M/1/N	**0.391**	0.329	0.168
Mark Procs	**0.627**	0.006	0.001
Tutoring	**0.514**	0.002	0.008
Proy Final	0.102	0.233	0.138

Table 2. PCA square cosines. Values in bold correspond for each variable to the factor for which the squared cosine is the largest

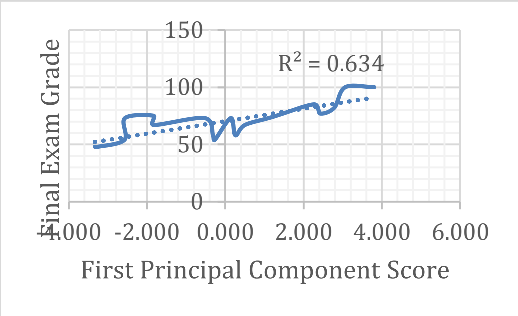

Figure 2: Relation between the first principal component and the final exam

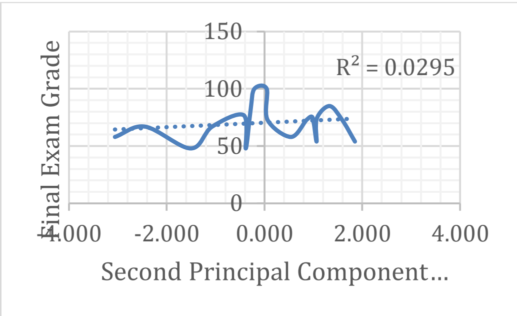

Figure 3: Relation between the second principal component (associated with a queuing homework) and the final exam

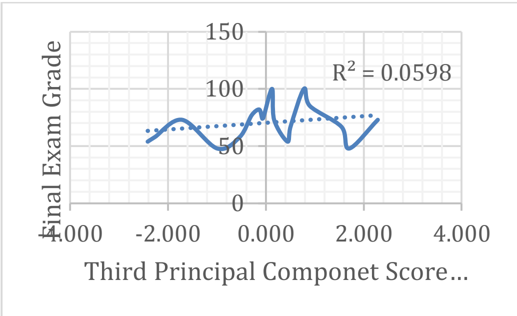

Figure 4: Relation between the third principal component (associated with a probability homework) and the final exam

Another important aspect to consider is, that if the first PC is associated with the most learning activities and has a positive correlation to the final exam score, then the first PC must be a discriminant to good and bad grades. That is, a low score on the first PC must mean a low score on the final grade. This is illustrated in case of the Stochastic Models course in figure 5. In this figure, a label “A” means a grade above or equal to 70% and a label “B” means a grade below 70%. There are 9 students right from the origin and of those, only 2 have “B” label, giving a 77.77% probability of getting a good score on the final grade if there is a good score on the first PC, and there are 9 students left from the origin of which 6 have “B” label, giving a probability of 66.67% of having a bad score on the final test if the score on the first PC is low. Also notice that in the upper right quadrant, the quadrant with high score in the first two PCs, all 3 students have label “A”, giving 100% probability of having a good final score if the score on both first PCs is high, whereas in the lower left quadrant, the quadrant with low scores on the first two PCs all 3 students have “B” label, giving a 100% probability of low score on the final grade if the score on the first two PCs is low.

It seems that the first PC is deterministic of the score on the final grade which raises an interesting point. Given this information, does this mean that we can finally eliminate final exams from university curricula? Although this seems likely, the main problem is that the variability presented in the correlation between the first PC. Recall that the R correlation is 0.634, meaning a significant correlation, but with high variance. Since the sample was rather small, we carried out the same statistical analysis with much larger sample, presented in the next section.

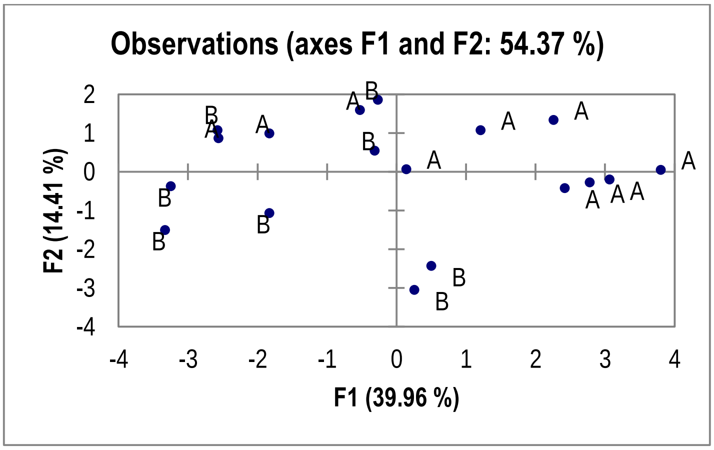

Figure 5: The first and the second principal components as discriminants of the final exam score.

4 Analysis of all Grades for Stochastic Models Course

For this new experiment, we gather all the known records of students for the Stochastic Models course, from 2008 to 2015, both spring and autumn periods. We intended to reduce the variability of the association between the first PC and the score on the final exam. It is well known that the variability of a sample mean in relation to the population mean changes with sample size. The average of a series of samples tends to the population mean, as the number of samples increases (Leon-Garcia, 2008). That is:

$$P\left[\lim_{x\to\infty}\frac{1}{n}\sum_{i=1}^{n}X_i=\mu\right]=1,$$

where

$$\mu=E\left[X_i\right],\ Var\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\frac{\sigma^2}{n},$$

and σ is the standard deviation of the population. Sample variance from the population mean decreases as sample size n grows to infinity.

In our case, 237 records were obtained. The significance of the sample size was estimated. This is easily done for normally distributed data. First, a normality test is carried out. The Kolmogorov-Smirnov (Yamane, 1967) test indicated normality on the data. See figure 6.

After determining that the data was approximately normally distributed, we can obtain the significance of the sample size by using table 3 (Yamane, 1967). Thus, 237 records, represent with 5% margin of error a population between 500 and 1,000 individuals, for example.

Now we proceeded to find if the variance of the relationship between the first PC and final grade score would diminish.

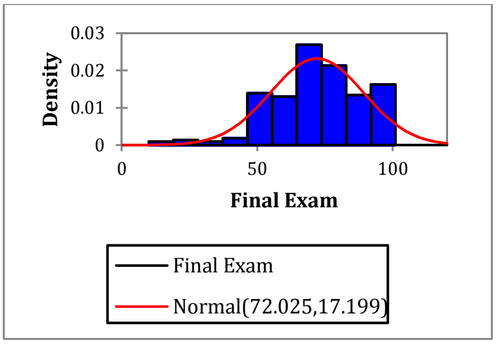

Figure 6. Histogram on normal pdf fit for all final exam records

Population	Margin of error		
	10%	5%	1%
100	50	80	99
500	81	218	476
1,000	88	278	906
10,000	96	370	4,900
100,000	96	383	8,763
+1,000,000	97	384	9,513

Table 3. Population, level of significance and sample sizes.

As the course changes activities in order to update and to adapt to the School calendar events and to correct errors or bias, the number of activities had to be reduced, as some activities are dropped with time and new ones are included. Not all the generations had the same activities. Also, some activities changed significantly during the course of 7 years. The activity count was reduced to only six. Two partial exams, and 4 home activities: Probability and random variables, two about queuing theory and one about Markov processes. The resulting square cosines are shown in table 4.

	F1	F2	F3	F4
1st Exam	0.242	**0.397**	0.002	0.035
2nd Exam	0.178	**0.506**	0.021	0.000
Prob rand var	0.429	0.000	0.048	**0.516**
M-M-1-∞	**0.591**	0.144	0.108	0.015
M-M-1-N	**0.651**	0.096	0.084	0.007
Mark proc	0.380	**0.016**	0.476	0.115

Table 4. Square cosines for all Stochastic Models courses

This time, exams were associated to their own PC, queuing theory activities also have their own PC, and just as in the Jan-May 2015 course, probability and random variables as well as Markov processes activities are on their own PC. That is not the expected result.

Nevertheless, the question still stands: Do these PCs have a positive correlation with the score of the final exam? That question is answered using figures 7, 8 and 9 (fourth PC, probability homework not shown).

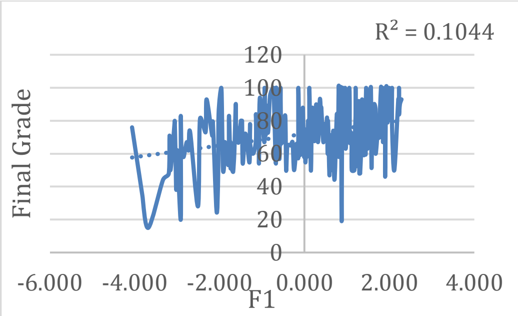

Figure 7. Final exam score against the first PC, associated with queuing theory

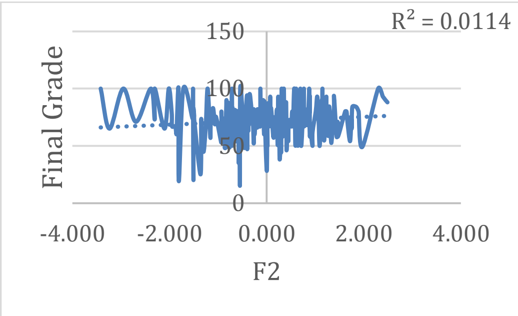

Figure 8. Final exam score against the second PC, associated with partial exams

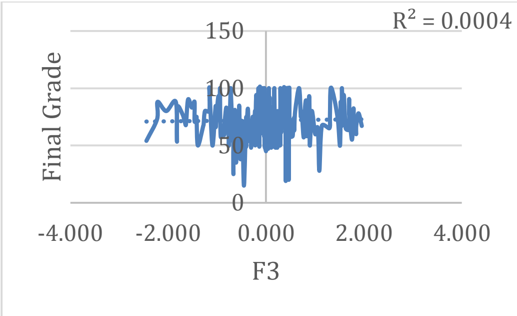

Figure 9. Final exam score against the third PC, associated with Markovian processes

The answer is yes, there is correlation between each one of the first four PCs exists. And this correlation is stronger with PC1, as shown in figure 6, and less strong with PCs two and three (and fourth not shown), as seen in figures 8 and 9.

Unfortunately, the R statistic of only 10% for the first PC indicates that the variability actually increased, not diminished. This is a reflection of the changes in the importance students give to the learning activities and of the fact that not all learning activities are included, and in some periods other learning activities not included in the analysis may hold higher importance to the student. As the linear fit represents only a small percentage of the data and such fit presents too high variability, we must conclude that final exams cannot be eliminated on the grounds that school activities drive students to better performance on cumulative final exams.

5 Conclusions

In order to determine the effect that carrying out learning activities in engineering education on the score on the final cumulative exam, we must first assess if the evaluation system is valid, reliable, recognisable and realistic, indicating that students understand what is to be evaluated, how is going to be evaluated and that students are actually evaluated in what they have learned. If the evaluation system complies with those characteristics then the question of having or not having final cumulative exams can be addressed.

We have shown that in a correctly designed and evaluated course, using PCA of all learning activities, the first PC will be equivalent to student's global work. Also, that the most desirable result is for the first PC to be associated to all, or most learning activities. Nevertheless, the nature of the activities and the attention and importance students give to some of them will make them stand associated to their own principal component, but keeping a positive correlation with final exam score.

Nevertheless, this correlation weakens as the sample size grows, indicating changes in the activities themselves and the importance students attribute to them. Also, as some activities are not included in the analysis because they are not common to all analysed periods, a lot of the explained variability can be hidden in such not included or accounted for activities as students may have given significant importance to them.

We therefore concluded, that with current evidence, and taking into consideration only cognitive factors, the elimination of cumulative written final exams cannot be recommended.

References

Alexandrov, N. S., & Ramirez-Velarde, R. V. (2013). The Integrated Learning Processes, Metacognition, and Collaborative Learning. In *Technological Advances in Interactive Collaborative Learning* (pp. 145-161). Boca Raton, USA: CRC Press.

Harvard, J. (2010, July-August). Bye-bye, Blue Books? *Harvard Magazine*.

Kirkpatrick, D. L. (1996). *Evaluating Training Programs: The four levels.* Berrett Koehler.

Leon-Garcia, A. (2008). *Probability, Statistics, and Random Processes for Electrical Engineering.* New York: Pearson.

Maltese, A., Tai, R., & Fan, X. (2012). When is Homework Worth the Time?: Evaluating the Association Between Homework and Achievement in High School Science and Math. *The High School Journal*, *96*(1), 52-72.

Myers, C., & Myers, S. (2006). Assessing Assessment: The Effects of Two Exam Formats on Course Achievement and Evaluation. *Innovative Higher Education, 31*(4), 227-236.

Pointek, M. (2008). Best practices for designing and grading exams. *Occacional Paper, 24*.

Ramirez-Velarde, R., Perez-Cazares, R., Alexandrov, N., & Garcia-Rueda, J. (2014). Education 2.0: Student Generated Learning Materials through Collaborative Work. *Procedia Computer Science*(291), 1835-1845.

Ramirez-Velarde, R., Sanhueza-Olave, M., Alexandrov, N., & De Marcos-Ortega, L. (2015). Do learning activities matter? *Interactive Collaborative and Blended Learning (ICBL), 2015 International Conference on* (pp. 76-82). Reykiavik: IEEE.

Svinicki, M. (1999). Evaluating and grading students. In U. o. Texas (Ed.), *Teachers and students: A sourcebook for UT-Austin faculty* (pp. 1-14). Austin: Center for Teaching Effectiveness.

Wiggins, G. (2014). *final exams vs. projects – nope, false dichotomy: a practical start to the blog year*. Retrieved February 17, 2016, from https://grantwiggins.wordpress.com/2014/01/01/final-exams-vs-projects-nope-false-dichotomy-a-practical-start-to-the-blog-year/

Wiggins, G. P. (1998). Educative assessment: Designing assessments to inform and improve student performance. San Francisco: Jossey-Bass Publishers.

Yamane, T. (1967). *Statistics: An Introductory Analysis* (2nd ed.). New York: Harper and Row.