The opinions of students on grades and grading systems, and how they relate to current assessment

Iiro Pyhäjärvi
682285A Bachelor's Seminar and Thesis
English Philology
Faculty of Humanities
University of Oulu
Spring 2016
Table of Contents:

1. Introduction ............................................. 1
2. Methodology and data ..................................... 2
3. Theory .................................................. 3
   3.1 The Common European Framework for language competence ................................ 3
   3.2 The how and why of grading .......................... 5
   3.3 The effects of grading ............................... 8
4. Analysis ................................................... 9
   4.1 Students in favour of numbered grades ............... 9
   4.2 Students against numbered grades .................... 11
   4.3 Improvements of grading system ...................... 14
   4.4 How the students’ views relate to the Common European Framework and theories of education .................................... 17
5. Conclusions ............................................... 19
   References ............................................... 21
1. Introduction

The goal of this thesis is to discuss whether or not language students of the University of Oulu feel grades to be important, and whether their views are compatible with current methods of evaluating students, and ask if indeed grading is even a necessary part of teaching. Grades are a source of excitement and frustration for students, and they have their proponents and opponents. The students’ views on them diverge rather drastically from one another. As a language student and currently completing my teacher training, the importance of grading students is of special interest to me, and considering how large a role grades play in the lives of students, their effect and effectiveness should be studied. My intention is to continue this research in my Master’s Thesis, hopefully gaining a more in-depth understanding of the issues present.

The data that this paper will examine comes from a survey conducted in the fall of 2015 as a part of the author’s pedagogical studies. The survey was aimed at students of the University of Oulu, and consisted of open ended questions regarding their grades in upper secondary education and university, the possible changes to the grades that happened in between, and the students’ reactions towards those possible changes. The survey also contained a question about the students’ general attitude towards grades. It is this question that this thesis concerns itself with.

The answers will be analysed through qualitative content analysis. This method allows the interpretation of different kinds of texts, focusing on the content and meaning of the text instead of studying syntactic structures and different ways of conveying a message. By closely reading the data, and then restructuring it into critical new interpretations, one can infer completely new information from the data set.

Once the answers have been analysed, they will be contrasted with the theoretical background of this paper. One cornerstone of theory is the European Framework for language competence. It tells teachers what skills students should possess at certain levels of language competence. This is taken into account in the national curriculum, and can be used in assessing students, especially in their final grades. University curriculums will also be examined. Basic principles of evaluation will be discussed as well, as presented by Atjonen (2007). Studies on the effects of grading in students by Veit (1979), Poorthuis (2015), and Gordon (2010) will be
examined, and the results of the questionnaire compared to them. A sociological view by Meighan and Siraj-Blatchford (1997) on teaching and assessing will also be looked at, and different theories of learning quickly addressed. The results show a tendency towards one type of teaching and assessing students, even though the subject of grades divides the students.

2. Methodology and data

In this section I will tell how I intend to analyse the data I have collected with the methods I have chosen, and why I have chosen those methods. I will also explain the method of data gathering, the chosen questionnaire participants, and the parts of said questionnaire that are relevant to this study.

As the focus of this thesis is in interpreting texts – the survey results – the method must be qualitative. In order to draw meaning from text, it has to be interpreted. Mere counting of words and grammatical structures is not enough, or even capable of doing this, so quantitative methods are disqualified right away. Qualitative analysis can be focused on either form, or content. The latter is the one this paper is interested in, so content analysis has been chosen as the method of interpretation.

Content analysis starts with the text. The text is assumed to be meaningful, as it is produced to be read and understood by others. Texts have no objective meaning in themselves, and they are thought not to exist independent of the reader. In the same way, texts do not have just a single meaning. They can be made for some specific purpose, but when read from another perspective, new information can be learned. The information and meanings that the reader infers from the texts also do not need be shared with other readers. In fact, different interpretations arise necessarily from different perspectives. Although this might make it look like no valid interpretations can be made, it is the context of the analysis that defines the meanings that can be inferred from it. The analyst has to construct a narrative in which the text makes sense and answers the question the researcher has set out to examine. Finally, the analysis has to be more thorough, more explicit and more verifiable than the interpretation of an uninformed reader. (Krippendorf 2004: pp. 21-25.)

The data that will be analysed was gathered in the fall of 2015 as a part of a survey regarding the change in students’ grades between upper secondary education and university, and the
students’ reactions towards those changes. The survey was conducted anonymously through the website FluidSurveys.com. The link to the survey was sent to the students through their major subject’s e-mail list. The survey consisted of twelve open ended questions, starting with the student’s major subject and their grades, and ending with a question about their general attitude towards grades and grading. This question is the one this study examines more closely. The question was “what do you think of grades and of giving them in general?” The answers range from short and concise to longer and more elaborate. If relevant information can be found from the other questions, they may be used as well, but the main focus is on this one question and the answers it got. As the survey was done in Finnish, the answers are in Finnish as well. They will be translated to English where necessary.

Originally the people questioned consisted of students of history and geography in addition to language students of English, German, Swedish, and Finnish. For this study, the history and geography students have been removed from the data, leaving only the language students. As the interest of this paper is in the opinions of language students, the exclusion of the non-language students can be justified. Even with the removal of these students, the data still has answers from 39 students. The students will be referred to here on out by the page number [#] they can be found in the data file, as there is only answers from one person per page, and it is therefore a natural way to refer to them.

3. Theory

In this section I will show relevant theoretical background for the thesis, and demonstrate why it is relevant. I will start by explaining what European Framework for language competence is, then explain how grading is done and why. I will also look at some studies on the effect of grades on students.

3.1 The Common European Framework for language competence

*The Common European Framework of Reference for Languages* – CEF – is a multinational guideline created by the Council of Europe for the purpose of providing a table of requirements for certain levels of competence, and a unified means of measuring the language
The CEF defines the general competence of language through knowledge, skills, and existential competence (CEF, p. 11). Knowledge refers to the knowledge gained from formal learning and experience, and skill or know-how refers to the ability to use knowledge correctly. Existential competence consists of individual characteristics and personality traits, including things such as how willing one is to engage with others in social interaction. The CEF also mentions the person’s ability to learn, which is something that comprises of all the above mentioned, and must be taken into account when teaching, when planning what to teach, and especially in how to teach. (CEF, pp. 11-13.)

The criteria for the levels of competence chosen for the CEF should meet the following criteria: The results must be able to be generalizable from different contexts. The scale should be equally applicable to school context and adult learning. At the same time, the scale should be relevant to different contexts, and usable for the function they are used in. The scale also needs to be grounded in theories of language competence, while remaining user-friendly and readable for non-academic users of the CEF. At this time, available theory is not adequate enough to provide a solid basis for the description of the scale (CEF, p. 21). This remains as one the most glaring issues of the CEF. The measurements of the scale and the different levels on it should be objectively determined to be based on theory of measurement. The levels should also be able to show progression from one level to another, and be distinct enough to
separate. The CEF claims to have met these criteria through a combination of intuitive, qualitative and quantitative methods. Existing scales were first analysed, then intuitively edited and new descriptors were formulated. Qualitative methods were used to make sure teachers understood and could relate to the descriptors. Finally, the chosen descriptors were scaled using quantitative methods. The accuracy of the scaling has been monitored in studies that followed. (CEF, pp. 21-22.)

There are six levels in the CEF, starting from A, going to B, and finishing in C. Each of these levels is also divided in two sublevels, one (1) and two (2), A1 being the lowest level, and C2 the highest. Each level has a description of the required knowledge and skills that must be attained in order to reach it. The CEF provides several tables for the levels, from more global and basic to more detailed and partitioned. Different aspects to measure include comprehension, written, and spoken language, as well as qualitative aspects of range, accuracy, fluency, interaction, and coherence. These tables can be used both to evaluate a student, and for the student to self-evaluate. In addition to table-form, the CEF also provides a branching tree-form, which helps with localizing, as it allows easier insertion of new breaking points between the levels. (CEF, pp. 22-33.)

3.2 The how and why of grading

The basis and justifications for teaching in general can be found in the national curriculum for primary school (Opetussuunnitelma - OPS). It answers the question of why children are being taught, and gives basic guidelines on how they should be taught. It also includes the fundamentals of assessing students. Elementary schools and secondary education have their respective national curriculums, from which they can structure their own, local curriculums. Universities do not have a national curriculum, but instead draft their own, which includes the structure of the degrees, the goals of learning for the degrees, and the content for each study module. The curriculums of Finnish universities are in line with the European Credit Transfer System, ECTS, which oversees that study credits are transferrable in higher education across Europe, but leaves the quality assurance for the institutions themselves (ECTS 2016). Thus the assessing of the students’ learning is on the shoulders of the course teacher.
In primary and secondary education the role of reviewing students is to encourage further learning in them (OPS 2016). This fundamental role of evaluation has also been written in the education laws of both primary and secondary education (Perusopetuslaki 628/1998, Lukiolaki 629/1998). The national curriculum gives teachers rather detailed goal posts for each grade in their own subjects (OPS 2016). In universities’ curriculums, no such things exist. The study guides give vague objectives to reach for each course, and the principles of grading are limited to simply listing the numerical grades, and their meanings. Even though this study is limited to university students’ views on grading, it is still good to contrast the differences in lower and higher education’s evaluation standards, especially given that the more transparent evaluation is the one the students are more familiar with, purely for reasons of spending more time in lower education. There is also no reason to assume that the principles of grading in universities would differ radically from that of primary and secondary education.

In her book *Hyvä, paha arviointi* [Good, Bad Reviewing] (2007) Päivi Atjonen lists several, ethically based values for good assessment of students. These include values such as fairness, validity, reliability, transparency, exactingness, ability to motivate, and the chance of showing one’s know-how. The principle that should most guide evaluators is to review others as they would wish to be reviewed themselves. In addition to this, Atjonen lists five other principles that Newman and Brown (1996) have compiled from the works of Karen Kitchener. First is respect of student’s autonomy, which means that while the reviewer has the power to determine the values of assessment, the opinions of the student should be taken into account when deciding what those values are. Second is the principle of avoiding harm. Needless harsh reviewing should be avoided, and negative critique should be presented in a constructive way. Third principle is the principle of doing good, meaning that the teacher should consider whose best they are reviewing for, the student’s or for some administrator in charge of funding. Fourth principle is fairness and justice, the equal treatment of students, and the taking of their individual needs into consideration. Fifth principle is the principle of reliability. The students should be able to trust the teacher’s capabilities in evaluating them, and that the grounds for evaluation do not change suddenly. These principles should all be followed when evaluating students. (Atjonen 2007, pp. 36-49.)

Meighan and Siraj-Blatchford (1997, pp. 164-173) define assessment as the act of collecting information on which the judgement of learning experience is based on. Assessment is usually
done by the teacher towards the student, but other forms of assessment also exists, such as students assessing each other, or inspectors assessing teachers. What is being assessed also changes according to who is doing the assessing. Inspectors, examiners and teachers stress the importance of the end-product, the student getting the right answer rather than the way it was achieved by the student. This view of what is important in assessing is absorbed by the students, who begin to put more effort into finding out what the right answers are, instead of actually trying to learn the subject matter. Meighan and Siraj-Blatchford report of a study by Becker (1968), where medical students abandoned the practice of learning to think independently and critically to instead try to glean cues to get the best grades from teachers. Also a common feature is the students’ view that discussion and debate are not real learning, and that dictation is the correct way of doing things, and recalling said dictation is the proper way of assessing the students’ knowledge and worthiness of a certificate. This ties well with the idea of a hidden curriculum, which conveys to the students how education should be done not by saying it aloud, but by being the way education is handled (Meighan & Siraj-Blatchford 1997, pp. 65-75). Hidden curriculum includes the practices that govern how knowledge and culture is transferred to the next generation, and in schools teaches students to respect authority, to passively accept instead of actively criticize, to think that feelings are inconsequential in education, and that there is always a single right answer to a question, among other things. When exams are the only form of assessment, students begin to think that exams are the only possible way of assessing students. Coming back to the learning experience being assessed, Meighan and Siraj-Blatchford list three different ideologies on how students learn. First is the transmission theory, where knowledge is something that the teachers transfers to the students, and assessment is done by evaluating how well the transmission of knowledge occurred. The second ideology claims that knowledge is passed in a dialogue, where the student’s understanding is reshaped in small increments in interaction with the teacher, and assessment is based on how well certain criteria have been met. Conversely, the third ideology states that learning is done by the learner alone, and the role of the teacher is to facilitate this learning by providing the learner with methods of acquiring knowledge, and assessment transfers from the teacher to the student as they learn to affect their own learning. (Meighan & Siraj-Blatchford 1997, p. 72.)
3.3 Effects of grading in students

On the surface level, the grades that students receive communicates the competence and achieved course objectives of the students. However, grades have additional, emotional value to the students. Students react differently to different grades, and those reactions can be a determining factor in their future grades. Veit (1979) points out that negative or lower scores in written papers do not automatically motivate the students to try to improve their writing, but leads to an avoidance behaviour, where the students avoid any risk taking, and try to minimize their errors. Veit also relates his observation that adding a grade to a paper changes the way students approach teachers about the papers: Without the grade, students are more likely to ask the teacher to elucidate their remarks, whereas adding the grade makes the students more likely to contest the grade, using any positive criticism as a proof against the grade. In Veit’s opinion, the grade distracts the students from the more important lessons they could learn from the teacher’s critique (Veit, 1979). A more recent study indicates a link between the perceived fairness of evaluation and the teaching and grading practices (Gordon, 2010). Students perceive the grades to be more just when the teaching practices are more focused on performing well on exams, as opposed to grading practices that might artificially inflate the grade. For example, when students are given the opportunity to discard one grade, it does not necessarily lead to better results, and might in fact cause the students to completely disregard that one grade, seeing as how it does not affect the overall grade. In general, such grading methods cause students to be less incentivized, and to think their grades less meaningful (Gordon, 2010).

A 2015 study on the effect of report card grades was conducted in Netherlands, investigating a possible connection between grades and behavioural and emotional engagement (Poorthuis, 2015). The study aimed to establish a link between positive and negative affect and motivation. Although the study was focused on elementary school students, the results can in the very least be contrasted with the questionnaire results, if not directly correlated. The study found a link between high grades and increased emotional and behavioural engagement, and also gave insight into why some students become disengaged from their studies. One of the findings of the study was the fact that passiveness and disinterested do not necessarily stem from negative feelings, but from lack of positive feelings. If a student receives no encouragement, or they are not offered any incentives to do better, they might feel that there
is nothing to achieve by increasing their efforts. This connection between emotion and the physical or mental effect it has is called affect (Wetherell 2012, p. 2). Affect theory postulates that affect is something that is always on, and is influenced by social actions. Sometimes the effect is immediate and intense, such as in panic attack, and sometimes more long term. In the Poorthuis’ study, the lack of positive affect over a long period of time, i.e. lack of encouragement or good grades, results in either negative affect, feelings of inadequacy or such, or no affect, where there is no motivation for either avoidance or validation.

4. Analysis

In this section I will analyse the data I have collected, and contrast it to the theoretical background of the thesis. The students’ answers will be referred to by the page number they appear in the collected data file. Translations will be provided where necessary. Although this paper examines only one question from the survey more closely, answers from others can also be brought up if they provide relevant insight into the matter. The opposing themes of students in favour of numbered grades and those against them was the most prominent feature found in the data, so it was given the most attention.

4.1 Students in favour of numbered grades

Many of the students who responded to the survey felt that numerical grading is a positive thing, and that it is something that should be kept. For many, such as students [1] and [10], numerical grades represent a measure of one’s own knowledge, a concrete proof of successful learning: “Otherwise I would not be able to get a clear picture of what level my know-how is” [10]. This, of course, assumes that the reviewer is grading the student fairly, and follows a guideline similar to what Atjonen (2007) outlines in her book. If no such guidelines are followed, and the grading is done with little regard towards the students best interests, then naturally the grade might not reflect the real skills of the student. It is interesting to note that of all the students who mentioned the grades’ role as a measure of skill, only one of them mentioned the possible invalidity of the grading process:
Grades are on the other hand a way to measure how the study material has been internalized, but on the other hand their formation feels a bit arbitrary, albeit that courses do present some form of criteria for the grades [34].

This seems to imply that most of the students that value grades as a measure of one’s knowledge have an implicit trust in their reviewer’s capabilities and in their method of assessing. Whether or not this trust in the teachers is earned is not within the extent of this study. One student also argues that possible learning disabilities and other detriments to learning are easier to notice when there is a gauge for learning [1]. This student does not provide any background for this claim, and while people with learning disabilities might have lower grades, they alone cannot be the determiner of such a condition. The answer also does not make it clear whether this gauge would be internal, as in the student should notice the difficulties themselves, or whether it is external, whereupon it falls to the teacher to compare the students against one another.

Some of the students felt that numbered grades help upkeep standards. As one student suggests, they are “absolutely important, so that the easy way out could not be taken” [24]. This correlates directly with the principle of fairness that postulates that students should be treated equally (Atjonen, 2007). If it were possible to get the same results from less work, it would be unfair towards those who worked harder, and might lead to lessening of motivation for the hard workers, and eventually even lower their skills. What might follow is that if the better graded students decrease their effort, the skill level between the former highest grade receivers and current ones becomes wider, and the combined skill level of the workforce lowers. This is naturally detrimental to the national and local economy. Therefore, as one student writes, it is necessary to make the distinction between those who have not studied well and those who know the subject matter after the course has ended [33]. In keeping with the themes of fairness and work life, one student mentions that it is somewhat fair that the ones who do well in their studies have proof of their skills, and therefore have an edge when it comes to future employment [32]. This presupposes the idea that employers look at or even care what the numerical grades of the job seekers are, something that will be looked at more closely later in this paper. However, if the employer does care about whether one did well in their studies or not, then of course those who do well in their studies have the upper hand when compared to those who did not put as much effort into their studies.
One of the positive effects mentioned by the students is the perceived motivation that good grades give. Good grades “[...] make you cheerful and generates lots of self-confidence” [1]. It is impossible to deny the good feeling that good grades might give to a student, but whether it has any bearing on the meaningfulness of the grades is debatable. It is not the function of grades to make students feel good, but rather it is a side effect, and while the feelings of students should not be dismissed, neither should they be given too much value. Another students mentions that while the grades might not indicate anything about the student’s real skills, they do motivate [13]. This is in accordance with the Poorthuis’ (2015) study that found a link between high grades and emotional investment. If a student receives good grades, and they perceive the grading to have been just (Gordon, 2010), they are more likely to engage further in their studies, or in the least to stay on the same level of engagement. It is important that grades are not artificially inflated, as in for example the score of a test raised by a few points so the student reaches the next grade level, or the meaningfulness of the grade disappears. These kind of practices run the risk of students not applying themselves fully in their studies, if they know their grades can be raised simply by being pleasant to the teacher. This is something that teachers and students studying to become teachers should be aware of, so that they would not give better grades to those students they have a more close connection to.

4.2 Students against numbered grades

Not all of the students feel numbered grades to be the best way to measure one’s learning. One student thinks the evaluation methods of both secondary and tertiary education to be old-fashioned [2]. According to this student, current methods of reviewing only encourage temporary memorizing of facts instead of instilling an inner motivation. This student also points out that these methods do not always measure the student’s actual knowledge correctly: For example, an exam might contain questions from the area of the course that a student might not have internalized that well, while having mastered the rest of the course material. One could argue, however, that any exam should contain only the most important and relevant parts of the syllabus, and that if the exam does not contain such, then it is the fault of the teaching and exam preparing methods, not the fault of the assessment methods. If the student does not know what the most relevant information is, or even how to identify it
from the rest of the material, then it is a fault of studying methods. Either way, assessment methods cannot really be held responsible for these kinds of lapses in judgement skills of the teacher or the student. This same student [2] also criticises exams and their grades from the lack of interaction with the student. If all a student receives from their exam is a grade, they have no way of knowing what it was that they got right or wrong. According to this student, there is no way to have conversations with other students or the teacher about the questions, and no way of correcting their mistakes. This is a mistake on the student’s part. All the teachers have office hours, during which students can go and discuss their grades with them, and conversing with one’s fellow students requires only willing participants. What this student possibly means, is that there are no opportunities for discussion during the course, which is true. Most courses have strict time limits, and therefore must prioritize teaching. But to say that conversation is impossible altogether is just not true; it only takes personal effort to make it happen.

Another student thinks that grades are a “necessary evil in modern mass education” [8]. The student goes on to describe how the goal of modern teaching is not so much to prepare people for life and its challenges, but to give them a spurious proof of their skills. The goal of education, according to this student, is only to hand out diplomas, especially in higher education. As the funding of universities is partially based on the number of graduations, it is not that big of a leap to assume that universities have vested interests in having students graduate in time (Ministry of Education). This is a fairly cynical view of current Finnish education, and therefore should be taken with a grain of salt. If any university in Finland would start to mill out diplomas, its reputation would suffer, decreasing the number of applicants, and thus decreasing the amount of funding to the university. It is also unclear how numbered grades come into all of this, as higher grades do not affect universities funding (Ministry of Education). Only way they would be relevant is the fact that if a student’s grades are too low, they will not be able to continue their studies. In that sense this student has a point, but numbered grades are not the only way of measuring students’ knowledge, and to determine whether they will pass. More on this will be discussed later in this paper. Somewhat similarly to this student, another student explains their negative view of grades saying that a low grade is a pass in the same way that a high grade [7]. As pointed out before, this is not exactly true. While individual course can be passed this way, some study programs have average grade
requirements for getting into them, such as the faculty of political science in the University of Helsinki (Faculty of Political Science). A student who only barely passes their courses might be doing themselves harm by excluding certain future studying possibilities.

One theme raised by a few students is the possibility that grades do not reflect the actual knowledge of the students [13], [16]. Earlier in this paper the possibility of an exam having questions only from the portion of the course the student is not familiar with was discussed, and argued against, but this does not prove that an exam might not actually reflect the student’s actual knowledge: A student could simply be having a bad day, due to personal reasons, health related issues or just lack of sleep [6]. It is not impossible to think that reduced cognitive abilities for the aforementioned reasons could affect one’s grades. However, once again this is not so much the problem of the grading system, but rather lack of preparation on student’s behalf. Furthermore, exams can be retaken, as can entire courses as well. Another way that grades might not reflect the student’s actual skills is if the grading is based on something other than pure demonstration of knowledge, such as Gaussian normal distribution [30]. One student argues that the personal opinions of the teacher can affect the grade as well [16]. In addition to the teacher assigning grades based on their opinions of the student, teachers might also decide grades by comparing students against one another [16]. These are not issues with grades themselves, but more in their implementation. As one student puts it, “[the fact that] grades are imperfect does not mean they are useless” [12]. The implementation of grades might be flawed, but that is the fault of the reviewer, not the grades or their justifications. The only argument against numbered grades that cannot be counter-argued easily is that grades create pressure to do well, and create competition [30]. As Veit (1979) observes, the mere presence of grades change the way students approach teachers. Low grades also lead to an avoidance behaviour, where students stick to tried and true methods instead of trying to improve themselves. Veit’s suggestion for combating this was to remove grades from most of the papers, and offer more written or verbal critique. More on possible to ways to improve upon the grading system follows in the next section.
4.3 Improvements of grading system

So far, this paper has discussed the students’ opinions about numbered grades, but has yet to mention the alternative, the pass/fail system. In this system, instead of having numbers denote the skill level of a student, a simpler passed or failed system is used to evaluate them. Many students seem to prefer this system that in one student’s opinion does not place the students into an unequal position when applying for work [20]. As pointed out before in this paper, it is debatable whether prospective employers even ask to see the applicant’s grades, but if evaluation was done in a pass/fail system, this would become redundant, and in theory at least increase equal opportunities in employment. Another student provides a good counter argument by appealing to the fairness of providing students with proof of their superior skills [32]. It can be argued that pass/fail system offers no incentives to put more effort into studies, as students who previously would have received the highest grades are put on the same level with those who did just average. One of the solutions offered by the students is to increase the criteria for a passing assessment [26]. According to this student,

[...] the most important thing is to understand and internalize the main content of the course, rather than to get a grade of five and possibly forget everything in a month [26].

This echoes the thoughts of another student mentioned previously, who felt that grading encouraged this sort of temporary memorizing of the course material, and did not lead to an inner motivation [2]. It would of course be better if students actually learned the course content instead of just regurgitating it onto an essay paper. It does not bode well for the future of the students, if they cannot remember what it was that they spent years studying. One student does however provide an argument against this, saying that there are always subjects that do not interest them as much as other subjects in their area of study [19]. If a student feels that a certain course is not going to be useful to them in the future in any way, the requirements of pass/fail evaluation system might be too harsh. Then again, it cannot be said that universities are supposed to make it easy for students to graduate. One student suggests that university could be completed completely independent of any university mandated syllabi [8]. In this model the university’s only role would be to facilitate and give tools for the student’s own research, and the student’s graduation would be completely dependent on the quality of their research. Problem with this model is that it does not take into account that
most of the students do not continue into an academic researching career after their graduation, but seek employment elsewhere, usually in fields where academic research is not a part of the job description.

One of the students mentioned a wish for more feedback, instead of the teachers just providing the students with their grades [27]. As mentioned by another student referred to before, when a student receives only a grade, they do not know what it is based on, what they did right or wrong in their exam [2]. Counter-arguments for this were given earlier, as it only takes some personal effort on the student’s part to find out where they erred. However, as Veit [1979] observes, when a paper is graded, students tend to focus more on the grade, instead of focusing on the lesson being learned. In his own teaching, he omitted giving students grades during the course, providing the students only with written critique, and only grading the papers at the end of the course. He noticed the effect it had on both the learning of the students, but also the work load of the teacher. He was also able to discern a difference between positive and negative reinforcement: By handing students examples of good writing as opposed to bad, the students were more eager to take part in the class. In his model, he still provided a grade at the end of the course, but the constant feedback during the course helped his students learn the subject matter better. Whether this is possible in courses that have a hundred students in mass lectures listening to the teacher trying to get the material to the students in both concise and sufficient manner is questionable. Some courses have provided a lecture specifically for students to ask questions, but most of these are held before the final exam, and are therefore of little help when it comes to explaining the grade. Even a little explanation for the grade would help the students, even if it is just returning the exam with notes. This would increase the amount of work that the teachers have to do, and in universities where research is a large part of the teachers’ expected work, this might be too much to ask for.

Self-evaluation is a theme brought up by some of the students [14]. In addition to the grade given by the teacher, the students should be given the opportunity to evaluate themselves, to give themselves a grade so that the teacher would know how well the students think they knew the contents of the course. The student thinks that this would lessen any confusion on whether they have over- or underachieved during the course. In their autoethnographical article, McClam and Sevier (2010) recount their experience with an experimental way of
grading their students. The students, who were studying to become teachers, were told that they would be allowed to evaluate themselves, and that this would be their grade for the course. According to the article, this was met with confusion and resistance from students. After having been assessed by the teacher their entire lives, the thought that they could grade themselves was foreign to them. The students were conflicted by their desire to achieve good grades and by their honesty. Some of the teachers also felt conflicted, having lost their role as the grade-giver. This proved to be the end of the experiment, as the people in charge of the school were uncomfortable with the loss of the traditional role of the teacher, and had the experiment cancelled. Despite these fears by both the students and the teachers, the grades the students gave themselves were similar to the grades of earlier years, and none of the teachers felt that any of the students had given themselves worse or better grades than they deserved. The students had difficulties in evaluating their own knowledge, or asserting how they even knew their own knowledge. This tension manifested itself as anxiety, stress and even anger. So, while self-evaluation might be something that students want to do more, care must be taken not to disturb the power relations between teachers and students, as they seem to be deeply entrenched in both of the students and the teachers.

Two of the students suggested that art subjects and physical education should not be graded [6], [25]. A question then arises, how should non-theory subjects be evaluated? Of course, not all art subjects are completely devoid of theory, for example in music education there is musical theory, notes, scales etc. But in subjects such as arts and crafts, woodworking/handcraft, and physical education, the grade seems to be determined by the quality of the work or the fitness capabilities of the students. One student brings up the effort they have put into the course as something that should be taken into account [2]. For example, with effort taken into account in physical education, the highest grade would go to the ones who tried their hardest to improve themselves, instead of the ones who were already physically fit to start with. However, in arts and crafts and handcraft the quality of the work cannot be ignored so easily. A poorly built birdhouse is a poorly built birdhouse, regardless of the effort put into building it. In subjects that require crafting skills, the function of the crafted object cannot be left out of the equation. Similarly, in music one might know all about the theory, but if they cannot put that theory into practical use, i.e. play an instrument or create music themselves, a good grade might not be warranted. The idea of effort being a
determining factor in assessment also goes against the idea of grades as a measure of skill. One student’s full effort might be another student’s average effort. It is also hard to quantify effort. Again, what is one person’s considerable sacrifice of time and effort might not be that much to another person.

4.4 How the students’ views relate to the Common European Framework and theories of education

The Common European Framework for language competence is a very organized and structured way of assessing competence. As such, it requires that students be able to be divided into different levels based on their competence, and therefore forms a natural connection to numbered grades. Each level of competence can be easily linked to a numbered grade, making assessment of students easier, faster, and even fairer, in the sense that everyone is being judged by the same unified standards. Among the students questioned there was a divide between those who support numbered grades and see them as necessary, and those who would do away with them. The proponents of numbered grades would have no problem with the CEF, and would receive its clear basis for the grades with open arms. While some of the opponents think that numbered grades are acceptable in lower education, they do not belong in higher education, where a student’s future should not be decided by grades on a course they might have no interest in, but have to take as it is compulsory [19]. At the same time, fairness of assessment is something that many students find to be important, and this is one area where the CEF shows its strengths, as the levels of competence are clearly marked, and available for all. Combining the CEF with a pass/fail system is difficult, as it requires the cut-off point between pass and fail to be well defined, and the justification for the cut-off point being wherever it is has to be sound. One of the common complaints about the current method of assessment is its lack of transparency. The CEF would make self-evaluation easier, as the standards are available for all to see, and have well-articulated explanations for all of its competence levels. For the complaint about lack of teacher-student interaction the CEF however does nothing. The CEF gives teachers the means to better grade their students, but offers no advice on how to relay to the student what they have learned, and what they have not. Also, as the CEF concerns the learning of language from the very beginning, it is of little use to university students, who should already have almost mastered the language they study,
and whose studies start to focus more on the deeper structures of language and its use. Even more problematic is that a grade is a single number that encompasses all the different competences that the CEF measures, such as reading comprehension, spoken interaction, and range. If a student’s grade is an average of several distinct factors, it is up to the teacher to explain to the student what their grade means. This was something that one of the students explained in their answer, that the imperfection of grades does not make them useless, but rather requires the teacher and the student to come to an understanding of what the grade means [12]. This student also expresses their wish that more studies on the effect of grades were done, before any changes in the grading system would be implemented.

Many of the students who answered the questionnaire seem to be under the impression that exams are the only way of measuring their knowledge, or in the very least it is to them the most stereotypical way of assessing students. When speaking of grades, they talk about how well they did in their exams, not how well or often they participated in the class, or whether any other course work affected their grade. As Meighan and Siraj-Blatchford (1997) put it, the students have started to stress the importance of the end-product, as have their teachers done when grading them. The students have absorbed the hidden curriculum of the ways of teaching and evaluating, instinctively resorting to them when presented with a question about grades. Even when calls for changes are made, they are directed towards teaching methods, and quietly assume that a final exam or an essay is still the best way to assess them. Students that care about numbered grades are far more concerned about getting that good grade than in challenging their own views of what is important to assess, and how it should be assessed. Even the suggestions for self-evaluation are linked to be a part of a final exam. This is alarming, as some of the students who answered the questionnaire are either studying to become teachers themselves, or have already completed their pedagogical studies. If they cannot even think outside of the metaphorical hidden box, they run the risk of just perpetuating the outdated theory of information transmission, where the student passively absorbs knowledge, and then presents the teacher with memorized factoids, without ever having to demonstrate any true understanding of the subject. Of course, any radical changes in teaching and assessing systems cannot be made, as the risk of being wrong has the danger of jeopardizing the education of an entire generation at worst, but change starts with awareness of the issues. None of the students questioned here have showed any signs of knowing about other theories
or ideologies of education. Of course, it has to be kept in mind that the original questionnaire did not contain any questions directly asking about such things, and was more concerned with the personal feelings of changing grades and grades in general. Still, when any reference to teaching or assessing was made, it assumed the teacher-to-student model of knowledge transference, and never really questioned the teacher’s authority in assessing the students.

5. Conclusions

All in all, the students agree that assessment is a necessary part of education, but have differing opinions on how the grading should be done. Some students find the current and most common system of numbered grades to be an effective way of measuring their skills and knowledge, while others would prefer a system of pass and fail. As the pass/fail system is hard to reconcile with the Common European Framework, it cannot be said to be very compatible with it. The few students that do not seem to care about their grades, or have negative opinions about them, have problems with the entire education system, not just the evaluation part of it. What is interesting to note is that none of the students mention anything about other theories of learning or assessing, only using the stereotypical, teacher led model where students passively learn and are tested at the end of the course with a final exam. A large number of the courses in university do not even have a final exam, instead opting for a comprehensive essay to be completed on the student’s own time. This speaks of an internalized model of education, absorbed from primary and secondary education, where the courses do follow this model. It seems that these models are hard to let go of, and in the case of students of pedagogy, do not possibly register as just one of the different methods of educating, unless specifically prompted to talk about them. When it comes to the principles of assessment, fairness is a term that is used many times, as is mutual agreement between the teacher and the students about the requirements of the course. The students are aware that assessment methods are something that can be changed or influenced, but do not seem to make the jump to thinking that what is being assessed can also be changed, at least according to the data at hand. Further studies on this subject will be conducted, as the author intends to continue this paper in his Master’s Thesis, using a new, better composed, and more form fitting questionnaire, and a larger and deeper theoretical understanding of the underlying issues.
In hindsight, it is easy to see that the Common European Framework was given too much attention in the beginning, as the results do not have that much to say about it. This was due to a few reasons: One, the author’s misunderstanding on what the CEF really is, a tool devised to help teacher’s grade students, not an educational theory on assessment; Second, the mistake was unable to be corrected because of time constraints, and the author’s habit of writing from start to finish, instead of writing in easily manageable and correctable blocks. As the results started to take the study in another direction, it was already too late to start the massive rewrites required. Instead, I decided to push forward and do the best I could with what I had. This is definitely something to be learned from, and corrected in the Master’s Thesis.
Sources


