



GRADING AND LEARNING: VARIATIONS AND COMPLEXITY FROM 19TH CENTURY UNTIL NOWADAYS

Department of Teacher Education
Master's thesis

Author:
Olha Nikolaienko

Supervisor:
PhD Koen Veermans

19.06.2024
Turku

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service

Master's thesis

Subject: Education and Learning

Author(s): Olha Nikolaienko

Title: Grading and learning: variations and complexity from 19th century until nowadays

Supervisor: PhD Koen Veermans

Number of pages: 84 pages

Date: 19.06.2024

Grading is considered to be an integral part of various educational systems with an influence on students' performances and learning outcomes. This thesis examines the multidimensional nature of grading from both historical and contemporary perspectives with an emphasis on interconnection between grades and learning.

The comprehensive review of the grading practices from the late 19th century until current days reveals various researches in the field of education which attempt to find a solution to such issues as subjectivity in grading, standardisation, fairness and the ways grades influence student's motivation. Yet, those areas remain open for further exploration since no unified solutions has been found.

The findings of the study highlight the complexity of the grading issue and suggest that in order for the grades to be more fair and promote learning, multiple assessment tools should be utilised. One of the tools that students view as the most helpful is feedback, however, even though it has its benefits, the study emphasises the importance of acknowledging that feedback is a form of grading that similarly to the traditional grading can be misinterpreted.

Moreover, the study overviews the variations across and within the students in their perceptions of the fairness in grading additionally revealing a noticeable impact on the psychological well-being of the students.

Various aspects of grading are researched through the prism of different researches throughout the time as well as the questionnaire responses, that has been used as the data collection method for this thesis.

Key words: grading, grade, evaluation, assessment, subjectivity in grading, variation in perception, fairness, feedback, standards-based grading, motivation.

Table of contents

| | |
|---|-----------|
| 1. Introduction | 4 |
| 2. Chronological literature overview | 7 |
| 1800 — 1920 | 7 |
| 1921 — 1940 | 10 |
| 1941 — 1960 | 13 |
| 1961 — 1980 | 14 |
| 1981 — 2000 | 17 |
| 2001 — 2010 | 21 |
| 2011 — 2022 | 26 |
| Overview of the main grading themes | 38 |
| 3. Methodology | 42 |
| Objectives and Hypotheses | 42 |
| Data collection instrument | 43 |
| Participants | 46 |
| 4. Results | 47 |
| Students' understandings of the grade | 47 |
| Students' perceptions of grades and evaluation | 49 |
| Grading and learning and their interconnection | 53 |
| Psychological impact of grading | 57 |
| The complex nature of grading | 59 |
| 5. Conclusion | 64 |
| 6. Discussion | 68 |
| Further Improvements | 69 |
| Limitations | 71 |
| References | 73 |
| Appendix 1. Questionnaire on students' perceptions of grading | 79 |

1. Introduction

Grading has always been an integral part of any educational system, since it is an essential part of the school and university experience — students receive grades throughout their entire learning path and, clearly, grades have an impact on the overall learning process. The most important aspect is to understand what the purpose of the grade is and what it shows. Furthermore, the aim is to analyse it as a complex issue that has an impact on student's learning.

The debates concerning the grading has its roots in the end of the 19th century and probably even earlier. The research of that time conducted by Edgeworth (1888) points out that numerical grade is inaccurate due to the fact that it is the person who decides which amount of knowledge is sufficient for the particular grade. Furthermore, Edgeworth compares assigning numerical grades to weighting something without scales — it is possible to do only a rough estimate rather than measure accurately. The author of the research emphasises the complexity of the grading issue in 19th century, and now in the 21st century grading is still considered to be a complicated, complex issue and there is no unified understanding on how the grading should work and what the grade should show.

First of all, it is crucial to provide the current understanding of what the grade is. When consulting different dictionaries of the English language for the definition of the word “grade,” various aspects are included in the explanation. Thus, Merriam-Webster online dictionary defines grade as “a mark indicating a degree of accomplishment in school” (Grade. 2022. In Merriam-Webster.com. Retrieved July 1, 2022, from <https://www.merriam-webster.com/dictionary/grade>). Cambridge online dictionary formulates the definition of the grade as “a number or letter that shows how good someone's work or performance is” (Grade. 2022. In dictionary.cambridge.org. Retrieved July, 1, 2022 from <https://dictionary.cambridge.org/dictionary/english/grade>). Longman online dictionary of the contemporary English provides the following definition of the grade: “a mark that a student is given for their work or for an examination” (Grade. 2022. In [ldoceonline.com](https://www.ldoceonline.com). Retrieved July, 1, 2022, from <https://www.ldoceonline.com/dictionary/grade>).

The above-mentioned definitions are undoubtedly similar and yet they include different aspects of the grade. For instance, Merriam-Webster's definition is the broadest one since it describes that grade is given as an indication of accomplishment. Whilst Cambridge

dictionary shows that work and performance are separate entities, which should probably be graded separately. So, even these dictionary definitions show that the grade is a complex entity that cannot be described and understood in one way only. Furthermore, the definitions provide the broad explanation without giving a context in which grades are being used. So, the change in the grading practices is not only bound to the way of grading differently but it should also consider the context in which that grading occurs. Depending on the country, community, educational system this context may vary (Brookhart, 2017).

Historically, numerous approaches have attempted to explain the meaning of the grades and the ways in which the grading process should be organised. And in the historical overview part, I will investigate how different grading practices changed or transformed or whether they remained unchanged until nowadays. The aim is to investigate the grading throughout the time and highlight the most prominent themes that have been researched. The results of the questionnaire, that is used as a data collection instrument in this thesis, will be later analysed through the prism of those main topics and will be supplemented with students' perceptions and perspectives of those themes.

The aim of my thesis is to investigate the complex and multi-dimensional nature of the grades and their influence on the learning on the basis of theoretical material such as chronological overview of the grading practices starting from the 19th century until nowadays and their possible change throughout the time. Additionally, during the historical overview different approaches towards grading and learning will be presented as they give an opportunity to observe different perceptions of grades and show the examples of how the grading can be implemented in the learning process so that it supports the learning process or vice versa. Finally, I intend to collect students' experiences about grading and their perception of grades in general as well as whether they consider grades as a source of motivation and a valid criteria for their knowledge, what they consider to be more important grades or learning, and what the ways of improving the grading process so that it is fair and supports the learning process are.

The research questions that are intended to be answered in this thesis work are the following:

1. What are the ways to minimise the discrepancy in perception of grading between teachers and students?

2. Is student's perception of fairness influenced by whether the grade is higher or lower than they expected?
3. Do students prioritise grades over the learning? Is it the grade or the learning outcome that students perceive as a valid measure of their performance?
4. Do students consider grades to be an integral part of their learning and a source of motivation for learning?

The **novelty** of this thesis work lies in the chronological overview of the different researches focused on the grading practices as well as their impact on learning and further classification of the most prominent themes together with their reoccurrence throughout the centuries. The above-mentioned classification will provide an opportunity to observe density of the grading research in the chronological order and allow to discover whether those topics are present in the respondents' answers. All of these aspects will create a current image of the grading situation and possible ways for improvement and further research.

Grading has been a subject for discussion for over a century, there have been plenty of debates on which grading practice is better and more beneficial and yet until nowadays, there are still plenty of uncertainties regarding grading. Analysing both previous and current perceptions of the grading will allow to observe possible patterns and common features and have a broader perspective on the grading framework in general.

2. Chronological literature overview

This section is intended to analyse in chronological order the existing knowledge about the perception of grading as well as the correlation between grading and learning from different perspectives.

The existing knowledge of the grading and grading practices is presented in the chronological order for the purpose of showing how the understanding of grading was developing and whether it was changing throughout the time. The idea behind this is to show multiple viewpoints towards grades and grading throughout the centuries. Based on the relevant pieces of literature it will be possible to obtain the general understanding of the way the grades were perceived at the particular timespan and thus later compare with the further views outlining the benefits and drawbacks of them.

The main part of the literature overview will provide an extensive chronological overview of the researches and the final part will provide the schematic overview of the timeline as well as point out the most prominent themes. This schematic overview will present an opportunity to see the possible gaps and density of the researches as well as allow to observe the continuance of the researched topics. Further those themes will be linked with the obtained results from the questionnaire. The conclusion and discussion parts will present the summary of the views and their benefits and drawbacks.

All those aspects combined will provide an extensive overview of the situation with the grades and grading and their impact on learning process and possibly outline the ways for improvements of the grading.

1800 — 1920

Grading is a complex and often subjective matter due to the fact that it involves human judgement. Edgeworth (1888) states that numerical grading is inaccurate because it is a person who decides what amount of knowledge is enough for the particular grade. Even though the teacher is highly knowledgeable they still have their own views and thoughts on particular subject, which can influence their judgement during the grading process. Meyer (1908) also points out that grading cannot be simplified into just mechanical work that considers only the percentage of the right answers. The issue is more complicated since, for example oral examination, is harder to fit into the frame of the notions of right and wrong.

The distinction in this case is more vague and consequently, the grading criteria is becoming unclear as well. In Edgeworth's view (1908), assigning a grade can be compared to weighting something without using scales. It is suitable to do a rough estimate of the weight, but this method of weighting certainly lacks accuracy. Same situation can be observed with the grading. Teachers can only do a rough estimate of student's learning, but the grade should not be taken as the only measure of the student's knowledge. Starch and Elliot (1912) mentioned that "...marks are far less precise than the majority of teachers and pupils believe" (p.456). This statement was based on the analysis that resulted in considerable deviation of the points of the particular student by different teachers. This leads again to the idea that the lack of standardisation of grades results in teacher's subjectivity towards a particular work or student. Meyer (1908) states that there is an issue of undefined criterion for the specific grade meaning that there is a grading scale but the definition of what should be done to receive a particular grade is vague. According to the above-mentioned author, it is believed that when the teacher has a grading scale, they know how to differentiate grades and they know what level of knowledge corresponds to a particular grade. However, the issue of undefined criterion might lead to a situation where the teacher and students consider the grade to be more random rather than a reflection of hard work.

Moreover, during the process of grading the student's emotional state is not considered (Edgeworth, 1888). It is, however, an important element since an exam can be a stressful situation for a student that can result in misjudgement of the knowledge because the mistakes that were made are not caused by the absence of knowledge in a particular subject but in the increased stress level. This is another ambiguity that is connected to the grading. On the one hand, the more factors are included in the grade the more misinterpretations could be observed. On the other hand, the broader understanding of the student's work is collected the better understanding of the knowledge can be received. Starch and Elliot (1912) also note that grades can have an impact on the student's attitude towards education and life in general. The authors underpin the idea that lack of standardised grading can result in unfairness and inconsistency, which can lead to student's demotivation in learning and disbelief in the educational system in general. They also mention that the deviation in the grade for the same paper, which was given by different teachers can be considerable. According to their study the difference in grade can reach up to 35-40 points on the 100-point scale. That is one more

argument to the statement that the grade is a subjective matter that varies from teacher to teacher.

Starch (1913) raises the question of reliability of the assigned grades through the prism of the previous researches as well as the fact that different standards exist among different teachers and their overall methods of grading, different counties and different schools and the variation is brought into the grade by each of this factor. In his research 10 instructors were grading one pupil's paper. The table below (Figure 1) shows the results of this experiment. It is worth mentioning that in some cases, as for instance, Paper 9 the difference is such that the work is either success or failure (in this research the grade below 60 is considered to be a failure).

| Papers | Instructors | | | | | | | | | | Ave. | Mean Var. |
|--------|-------------|----|----|----|----|----|----|----|----|----|------|-----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | | |
| 1 | 85 | 84 | 87 | 90 | 89 | 80 | 83 | 81 | 86 | 86 | 85.1 | 2.5 |
| 2 | 77 | 78 | 77 | 85 | 76 | 82 | 77 | 81 | 86 | 86 | 80.5 | 3.5 |
| 3 | 74 | 76 | 77 | 80 | 83 | 84 | 86 | 85 | 80 | 79 | 80.4 | 3.3 |
| 4 | 65 | 63 | 61 | 25 | 40 | 60 | 50 | 49 | 56 | 49 | 51.8 | 9.2 |
| 5 | 68 | 80 | 77 | 87 | 78 | 88 | 80 | 79 | 79 | 79 | 79.5 | 2.9 |
| 6 | 94 | 85 | 92 | 92 | 97 | 77 | 84 | 83 | 89 | 88 | 88.1 | 4.7 |
| 7 | 88 | 88 | 94 | 92 | 93 | 85 | 81 | 85 | 88 | 88 | 89.2 | 2.6 |
| 8 | 80 | 82 | 72 | 84 | 86 | 83 | 80 | 85 | 78 | 75 | 80.5 | 3.5 |
| 9 | 70 | 68 | 67 | 55 | 58 | 65 | 70 | 75 | 80 | 78 | 68.5 | 6.0 |
| 10 | 93 | 90 | 84 | 97 | 95 | 83 | 87 | 83 | 85 | 84 | 88.1 | 4.5 |
| Av. | | | | | | | | | | | | 4.3 |

Figure 1. Grading experiment results by Starch (1913).

When analysing the results how the above-mentioned differences impact the reliability of grading in general the following statement was concluded: it is more reliable to grade students not so precise in the case when the grading scale is up to 100 meaning that instead of 100, 99, 98, 97, etc give 100, 95, 90, 85 etc. because it is almost impossible to notice such a small difference in the quality of paper, the bigger gap will allow to grade more accurately. In the case of A, B, C, D grading scale it was suggested to use "A+, A-, B+, B-, C+, C-, D+, D- and failure" (Starch, 1913, p.633) and that will bring more precision to the grades and consequently more reliability. That is one approach that Starch (1913) suggests as a solution to the reliability problem.

Although, the availability of researches of this time period is limited, judging from the ones that were mentioned above it is possible to observe some common features. First of all, all of the authors state that there is a problem of the lack of standardisation for the grading process and the subjectivity of this process due to teacher's personality with the specific views and feelings, which result in students' losing their motivation in learning as well as in education in general. Furthermore, the fact that the grade is one-sided i.e. student's personality and wellbeing is not considered was also seen as a problem even in the end of 19th century. Overall, the researches emphasise the importance of standardised grading but simultaneously the need for flexibility in the assessment.

1921 — 1940

The discussion around reliability of grades and their impact on student's learning continued later as well raising similar set of questions. For example, Sillberstein (1922) explores the impact of the teacher's personality on the grade assignation. The impression that a particular student made upon the teacher will lead to the situation where one examiner accepts the work, and the other does not. The article also shows the lack of uniformity among teacher grades. Teachers, who are considered to be experts in their field, were asked to grade the paper in English, where all student's identification was hidden. The task was to write an interesting letter about a hobby or a study. The difference in grades was significant — some teachers had given the highest grade, and some had given a half of the possible points. According to Silberstain (1922) the difference in grades might be affected by the way the teachers understand the notion of "interesting". Since people have difference personalities, their vision of what is interesting is likely to differ as well. So, it is probably worth considering eliminating questionable terms and notions such as interesting, exciting, etc as much as possible since, as it is visible form the research by Silberstein (1922), they interfere in the grading process by referring to teacher's personal understanding of the above-mentioned terms. It does not mean that by avoiding these terms the grading will become purely objective and unified. Moreover, the number of such terms is immense and thus it is almost impossible to eliminate them from the assignments. However, the teacher might give an explanation of the term to minimise misunderstanding and state the expectations to the work and to the

extent of “interestedness”. All of that might bring more clarity into the overall grading process.

Hulten (1925) emphasises the problem of subjective nature of grades even more by bringing the problem of inequality that can come together with the grading. The author supports the above-mentioned statement by the fact that based on the grades assigned by teachers it is decided whether the student will receive a place in the higher educational institution. Consequently, even the slightest inconsistency in teacher’s grading can have a direct impact on student’s future. When Hulten (1925) conducted an experiment with teachers grading pieces of students’ work several times during some time, the result was that 20% of the teachers changed their grades from “pass” to “fail” or vice versa. That is the fact that supports the thought that the grade is a subjective matter, and it is likely to vary not only among different teachers but also among one teacher but in the different time periods. It is worth mentioning, that based on the article by Hulten (1925), this discrepancy in grades is not the result of the teacher being unprofessional or lacking experience. The complication here lies in the fact that grading an essay (Hulten (1925) took students’ essays for researching consistency in grading) is the task which will most likely always be opinionated. This can be an underpinning point to the research by Silberstein (1922), where the letter about interesting hobby was graded and the inconsistency in grading was found as well. Additionally, the experiment conducted by Eells (1930) brings one more perspective to this issue. After testing the approach of the same teacher grading the same paper after a certain period of time and another teacher grading the previously graded paper, the following results were found. They showed that the higher judgement was demonstrated and there was no significant difference whether the same paper was regraded by the same teacher or by a different one. Either way, the reliability coefficient is low. And thus, the grade itself is not the embodiment of the true student’s knowledge. Consequently, it might be stated that “teachers’ marks are mere guesses, some good, some poor, some indifferent” (Hulten, 1925, p.54).

So, one suggested approach is the substitution of the essays by examinations with more strict standards, such as true-false and questions with multiple answers can bring more objectivity to the overall grading process. In spite of the fact that applying such methods will show students their progress more clearly, in practice, essays also reveal students’ thoughts and creative thinking in a way that no other examination can show. The solution might be that

essays should not be the only piece of work that determines students' final grade. Other, more objective, ways of examination, some of which were mentioned above, may be added. As a result, the final grade might still contain partial subjectivity, however, it will be minimised.

Crooks (1933) summarised different views on grading practices as well as their necessity. Some of the researches which Crooks use as references are impossible to find nowadays and thus, they can be analysed only through the prism of Crooks' vision. A number of authors of that time (Morrison, 1926, Odell, 1929, Crew, 1930, Hillbrand, 1930) suggested that the absence of grading might reveal intrinsic motivation and bring the focus to the learning and educational process instead of striving only for the highest grade. The debate around zero-grading can be found nowadays as well. But it is interesting to know that the discussion about that has its roots in the beginning of the 20th century. However, other researchers (Lund, 1929, Madsen, 1931) argue that comparison is inevitable in the classroom and that students will be compared either by the means of different grades or different feedback or some other methods. There still will be students who are more knowledgeable in a particular subject and who are less. The more important matter here is which emphasis is given to this comparative nature and how obvious it is.

The complicated issue of reliability of grades has also been brought up by Crooks (1933) because even with the number of experiments conducted in different settings, using different grading scales and graded by different teachers after a certain period of time or by the same teacher, there is still no unified explanation of the ways of how to make grading reliable. As long as there is a human being involved the judgemental element will always be present. However, the necessity of clearing the content of each grade has also been presented. Meaning that the grade should be at least communicated in order for the student to understand teacher's vision of the work.

It is worth mentioning that such factors as interest, motivation, attitude and effort have been a subject of debate in regards to whether they should influence the grade or should the grade be purely an indicator of academic achievement. But then again, if the teacher is the one who grades the paper, then the grade does not indicate only academic achievement but also a teacher's attitude towards the student and/or subject.

Crooks (1933) demonstrates different views on different grading scales such as percentage scale, ranking, ability grouping, use of normal and opposite normal curves. Each of them has

shown both positive and negative aspects and all of them have an issue with the reliability. For instance, on the one hand, percentage scale has been shown to be reliable enough however, on the other hand, the 5 percent difference between two papers is not considered to be significant. So, in such a case the scale should be designed in such a way that there is some gap between the percentages i.e. instead of having 100, 99, 98, 97... it would be more useful to have 100, 95, 90... .

The summary presented by Crooks (1933) has revealed that achievement is considered by most researches at that time to be the most logical ground for the grade. However, it is not a single-sided issue but a complex one, so it should be treated as such. The grounds for grading, standardisation of the tests, teacher training, grading clarification — all those aspects and even more should be addressed as a complex issue for the sake of grades being a valid representation of students' knowledge.

The researches of this time frame have their scope not only on the need for standardisation of the grading practices but they try to deepen their understanding of the reasons behind the particular grade, where human nature is considered as the one of the main sources of subjectivity in grade assignment and there is no unified vision on the solution to this problem. However, with more exploration and more visions on the grading, the better understanding of the complexity of the grading issue comes as well.

1941 — 1960

Swineford (1947) in his research attempts to discover correlation between the scores in arithmetics, verbal tests as well as sex, age and grades in English, history and reading in elementary school and the academic achievement in high school. It is believed that arithmetics tends to show the general intellectual level of the person and the verbal tests are meant to check the language capabilities. The author assumes that if there is a correlation between these two entities then there is a sense in grading students and if not, then grades are useless. Such an experiment is designed to show the reliability of teachers' grades. The results of this experiment demonstrated that teachers have a tendency to assign reliable grades to students based on correlations found. However, the fact that during this study such factors as age, sex and citizenship were included, makes those results more questionable since previous studies by, for instance, Crooks (1933) and Eels (1930) point that the inclusion of those

measures is bringing more judgement into the grade, making it less about academic achievement.

The study by Carter (1952) is aiming to find possible correlation of the sex of teacher and students and the assigned grade. There was a tendency noticed that girl students are getting higher grades despite the teacher's sex. Additionally, women teachers are assigning higher grades in general compared to the men teachers. Simultaneously, there was no striking difference in students' knowledge and that leads to the conclusion that the gender affects the way the work is graded. When the grades were compared to the general intelligence test, it was discovered that grades that were assigned by men teachers to student boys had the smallest deviation. Simultaneously, when comparing those two components in the group when women teachers were assigning grades to student girls this deviation was the biggest however not to the extent to make this difference significant. Even though no significant correlation was found during the study, the patterns were still found that the girl students were favoured more by both teachers' genders and consequently the grade contains not only the achievement and knowledge components but also the subjective element, as in this case students' gender.

The researches of this time frame are focused on even more details than the previous ones. These ones are aiming at exploring the grading further by the means of including gender, both students' and teachers' as well as nationality, age, etc and exploring the possible correlations and impact on the reliability of the grades. To which extent it is necessary to include these factors in the grade is still questionable because they do not show directly the knowledge and competence in a particular subject. However, both Swineford (1947) and Carter (1952) consider grades to be valid measure of student's achievement. They acknowledge the fact that the issue of subjectivity exists, but they emphasise that the standardised tests are the ones that measure student's achievement precisely enough.

1961 — 1980

The study by Holmes (1971) aimed to find correlation between students' evaluation of their own work and the evaluation of the teacher. The research revealed that students that were expecting higher grade showed more involvement in the learning process, were more motivated and more eager to ask questions. They described their teachers as reasonable and

adequate in assigning grades. However, the study did not find any correlations between lower grades expectations and students considering their teachers not professional or reasonable enough when the latter grade their works. The study also has not found any proofs as for the fact that those who expected higher grades were more intelligent. Some patterns were found but they were insufficient to consider them reliable enough. Nevertheless, the perspective that the author chose during his research is the one that was not studied well before and the fact that this one did not find any significant correlations is noteworthy because there might be a possibility that such a relation as students' expectations of a certain grades and their evaluation of teachers is not strong and does not interfere with the learning process crucially. However, in the study by Abrami, Dickens, Perry, and Leventhal (1980), where 264 undergraduate students in psychology were surveyed to evaluate teacher's effectiveness and the quality of instructions, the hypothesis was that the higher teacher standards are in grading, the lower students evaluate the quality of the instruction. The research concluded that student's perception of the fairness of the grading was of a higher value than teacher's actual standards. Consequently, there is a need of providing clear instructions and communicating the expectation of the course in order to provide clarity and feel of fairness of the grading process. This study shows that the correlation between students' expectations and the perception of teacher exists and thus it should be one of the aspects that is considered when discussing the issue of grading.

Further researches by Ebel (1974) and Collins and Nickel (1974) focus their studies more on the efficiency of different grading practices. The article by Ebel (1974) reveals the complexity of the grading issue which consists of lack of system for measuring knowledge and grade assignation. Moreover, the issue is further complicated when the educational practices start to emphasise the importance and uniqueness of each student. In the light of that any grading system is becoming too narrow to measure individuality as well. No standards were made to include that factor. Ebel (1974) examines different grading methods such as letter, numerical and pass/fail and explores their impact on the learning process as well as their strengths and weaknesses. The aim for all those explorations is not to eliminate grading but on the contrary, to emphasise its importance to the learning process. The main idea is to consider different evaluation methods in order to receive the most comprehensive view of the knowledge acquired by a particular student. So, the teachers have to understand the limitations of the

grading in general and try to use multiple ways to track student's achievement for the sake of receiving broader understanding of student's success.

Simultaneously, Collins and Nickel's (1974) study examined grading practices in institutions of higher education and found that grading practices varied considerably across institutions and academic disciplines. The study revealed that the majority of instructors used a grading system based on a combination of objective and subjective factors, with subjective factors playing a larger role in humanities courses than in science and math courses. For instance, the teachers of humanities rely majorly on the subjective assignments such as written tasks as well as class participation, where science and math teaches use test scores and lab results as a basis in their grade assignation. That is a natural process in view that humanities courses are more creative and require more personal involvement.

The study by Collins and Nickel (1974) reveals the issue of lack of standardisation of grading practices across the educational institutions, i.e. grading practices vary immensely between institutions in view of the fact that they do not have one standardised set of rules for doing that nation-wide, for example. The solution suggested by authors is to standardise grading practices more in higher-educational institutions as well as between academic disciplines and find the ways to bring more objectiveness to the humanities courses.

Both of these studies emphasise the issue of the lack of standardisation and the ambiguity of grades that appear due to this problem. Interestingly, they have similar view on how the insufficiency of standards in grading limits the understanding of student's success. In both of these studies the idea of applying various grading practices in order to achieve a broader perspective of student's knowledge is the one that worth considering.

Simon and Bellanca (1976) in their work "Degrading the Grading Myths: A Primer of Alternatives to Grades and Marks" also exploring the issue of standard grading and suggest the application of the alternative grading practices such as self-evaluation, narrative evaluation and providing formative feedback. In such a way the evaluation process gives more information to students on how to improve their performance and not limited just to the letter or numerical grade that can be construed in a different manner and can be easily misunderstood. The authors imply the idea that summative grades at the end of the term provide less information and interfere harmfully with the learning process because the emphasis is shifted from the knowledge acquisition and learning to achieving higher grade.

In conclusion, comparing this time frame with the previous one, the issue of subjecting in the grading process is remaining one of the most crucial problems that needs if not to be solved than to be minimised. The articles provide valuable insights on how to achieve that by communicating more clearly the expectations and requirements of the course as well as applying more grading methods for the purpose of acquiring more information of student's performance.

1981 — 2000

Marsh (1984, 1987) in his studies has been exploring the effectiveness and reliability of students' evaluations of university teaching. The dimensionality of students' evaluations can be viewed through the perspective of cognitive and affective factors. Cognitive factors include teachers' instructions and the clarity of them, organisation of the course in general and the feedback that was or was not provided for the assignments. Simultaneously, affective factors included teacher's enthusiasm and friendliness as well as sense of humour. Even though Marsh (1984, 1987) found that these evaluations are reliable enough and can be used as a source of acquiring information on how students perceive a particular teacher and also how effective used teaching methods are, it was mentioned that since the evaluation contains subjectivity that was stated as affective factors, these evaluations should not be taken as the only measure and other more complex methods of evaluation need to be used. Moreover, Marsh in his study "Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research" (1987) stated the need for the longitudinal study that might bring a broader understanding on the long-term impact of teachers' effectiveness.

The article by Jansow and Eison (1990) explored the correlation between grades and their influence on students and faculty. The authors argue that grades are the essential part of the educational process, but they acknowledge that at the same time they have a significant influence on students' motivation and behaviour. There is no arguing with the fact that grade have a psychological effect on students causing stress, anxiety and loss of motivation for studying overall. In the study Jansow and Eison (1990) explore a more collaborative approaches that might be used in order to mitigate the negative outcomes that are caused by grades. Providing meaningful feedback together with the grade might be a way for students to

reduce stress from the grades since it will not be just a number or letter, but it will be a source of improvement and understanding how to do the work better next time. Overall, the authors mention that holistic, more nuanced and complex approach needs to be used to improve the grading process so that it will be more beneficial for both teachers and students.

Winter (1993) in the article “Education or Grading?” views grading as such that contradicts the whole idea of education limiting the evaluation of knowledge only to some specific topics while the education should be more about becoming prepared for the future career mastering problem-solving skills as well as developing critical thinking. However, those are the areas that are not included in the grading process or included only slightly. So, the author argues that grades shift focus from the actual learning to competition and receiving higher grade and suggests to look for the ways of grades to be more informational and beneficial to the education and learning and less competitive. One alternative grading method that Winter (1993) suggests in his article is to use narrative evaluations i.e. detailed feedback for the work because this approach encourages students to actually learn and improve rather than just strive for the highest grade. Another alternative is the use of pass/fail grading scale. With such grading the evaluation is based on whether the student met a certain performance standard and it does not compare students by the way they perform a certain task but rather how they succeed in the course in general. The aim of these two grading approaches is to create a more collaborative and supportive learning environment.

Susan M. Brookhart has a great impact on the exploration of the grading issue and the impact of grades on the learning process. In her article “Teachers' grading practices: Meaning and values” (1993), she defines grades as an important and essential part of education, but she also acknowledges the complex nature of grading, which involves more than just assigning a number or letter grade to students' work. Moreover, grades have its undeniable impact on students and their perception of education, influencing their motivation and eagerness for learning. The article explores the aspects that might influence the grade assignment such as teacher's perception and attitude towards education, his/her beliefs, relations with the students etc. It is hard to estimate how big impact those aspects have on the grade since they have subjective nature, but Brookhart (1993) argues that they determine the way teacher views meaning and value of grades. Even though, Brookhart emphasises the importance of grades, she argues that grading is not a perfect system and in order for the grades to have meaning

and value it is crucial to realise that the function of grades is not only to evaluate but also to motivate, to provide support and feedback, to give a ground for improvement. The author agrees that grading comes with challenges, and it is the issue that requires constant further improvements and discussions in order to become more informative and eliminate misunderstanding that can come with it.

In her further research “Teachers' grading: Practice and theory” (1994) she discussed the development of fair grading system that can offer a field for improvement for the student as well as be perceived as reliable and valid. It is crucial that the aim of the grading is not comparison or ranking students, but it is first and foremost students' learning and students' improvements. Thus, teachers have to adjust their methods to make grades only one of the tools that the teacher has and a students can use for own development and not the final goal that students need to strive for.

Brookhart's research “Teaching about communicating assessment results and grading” (1999) is a logical continuation of her previous work with the focus on the importance of providing meaningful feedback together with the grade. In her opinion, the grade without proper explanation is meaningless since it does not provide any information that can be used for the improvement purposes. She argues that grades are not the tool that can be used to motivate students or can be based on purely subjective criteria. On the contrary, there should be a set of standards and goals that has to be, firstly, clearly communicated to students, and then used as a basis for evaluation. Brookhart (1999) offers practical solutions on how to improve communication of the assessment results by means of providing clear constructive feedback as well as communicating clear rubrics that are used in the evaluation process. Moreover, the researcher suggests that the implication of different evaluation formats and including students in assessment process as well as teachers' constant professional development are the crucial element of successful assessment results communication.

The study by Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995) stresses the importance of understanding teachers' assessment practices. The fact that there is an insufficiency of teacher training in terms of assessment, i.e. during teacher education programs, a little emphasis is placed on how to evaluate students' performance properly, how to use grades so that they support learning process and not interfere in it. Consequently, teachers rely a lot on their own experience and their own views during the grading process which leads to more

subjectivity and more misunderstanding between teachers and students. Moreover, without a proper understanding of different grading practices, teachers start to experiment deliberately with the practices without proper vision of how they align with the general instructions. Addressing these issues, according to Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995), and exploring the ways teacher education can be improved in terms of evaluation, can improve the quality of the assessment practices as well as the quality of learning outcomes.

Guskey (2000) critiques traditional grading practices which involve averaging grades and states that it might be detrimental for the overall learning process since it does not demonstrate the real students' knowledge in a particular subject. As an alternative the author suggests assessing students' performance based on the mastery in a specific learning goal. In such a way there will be a clear vision of what the student is knowledgeable about. Guskey argues that it is crucial to provide a clear and informative feedback along with the grade. Additionally, self-reflection and self-assessment as well as peer assessment are beneficial complementing grading methods since they include student in the process of assessment and thus a better communication is formed between teacher and student in terms of evaluation.

McMillan and Nash (2000) examine in their articles the factors that influence the process of decision-making when assigning grades to students. Among the key elements the authors highlight teachers' beliefs and perceptions regarding grading, teachers' own relevant knowledge in the field of assessment, school's policies and student behaviour and personal characteristics. Similarly to Guskey (2000), McMillan and Nash (2000) suggest teachers use more systematic and intentional approach in the process of evaluation and also include students in the process of grading through the medium of self-evaluation. Moreover, precise and clear instructions have to be given before the assignment and a constructive feedback is provided after completion of the assignment. The main idea is to use grading as a mean of promoting learning rather than just a tool for measuring knowledge.

The researches of this time period are focused on receiving a broader understanding of the correlation between grading and education. Authors are researching not only the purposes and necessity of grades but also the impact of grades on the psychological aspects that are closely connected with grading. Moreover, the researches are limited not only to those topics but also how grading influences the perception of teacher. The importance of feedback was stated in a number of articles that leads to the conclusion that it is an important aspect that needs to be

considered properly and explored further. And numerous articles of this time period view constructive feedback as a mean for improvement of the evaluation process as well as the way to make grading more fair and more reliable.

2001 — 2010

Guskey (2002) continues researching grading practices analysing traditional ones and exploring new, alternative ways that might be more beneficial for the learning process. An interesting perspective that the author takes is that he examines perspectives of teachers, students and also parents. Guskey (2002) highlights the importance of communicating assessment results clearly between all those three parties leaving a space for discussion and also communicating the expectations and the goals before the assignment in order to avoid misunderstandings. Moreover, meaningful feedback about progress and the areas for further improvement is a necessary tool in evaluation process. However, Guskey (2002) also emphasises that not only feedback but also student's own reflection on the work is important for further improvement. The author suggests that students and parents should be included, to some extent, in the evaluation process as it will bring more transparency and clarity to the whole process.

In the article by Smith (2003) the issue of reliability of classroom assessment is discussed. The author examines different factors that have an impact on grading and argues that a more complex and nuanced approach towards classroom assessment should be implied. Smith (2003) mentioned the judgemental role of the teacher as well as the subjectivity that comes along, and suggests usage of more clear assessment criteria and standards for the sake of making grading more about learning rather than judging. He suggests that multiple assessment methods should be used, the assurance that the clear instructions are given to students and also the feedback is provided to the students are the components that when applied can lead to a more reliable grading.

The issue of reliability of grading has also been raised and researched by Woodruff and Ziomek (2004). They explored this issue through the prism of investigating the inflation of the grades over the period of 12 years. The study showed that the percentage of higher grades increased significantly while the percentage of lower grades decreased. There are plenty of factors such as changing of teaching practices, student's perception of studying and the

increase in their motivation contributed to this inflation. However, the authors are cautious that the increase in higher grades does not necessarily mean that the academic achievement increased and consequently the reliability of these higher grades in relation to actual knowledge might be questionable.

Allen (2005) addresses the criticism towards the subjectivity and the potential inflation of grades and argues that these issues might be solved or at least minimised by the virtue of additional teaching training and standardisation of the grading practices. He mentions that grades are a valid measure of students' academic achievement and an important component in communicating information about students' performance to teachers, students and parents and it conveys a clear summary of students' performance and learning. Allen's view on grades is quite broad since he provides benefits as well as acknowledges the limitations that come together with grades such as inclusion and influence of non-academic factors.

Meuer (2006) in the article "Cognitive dissonance or revenge? Student grades and course evaluations" examines the correlation between grades and evaluation of a certain course. The results were that students that receive lower grades are more prone to report negative evaluations of their instructors which means that receiving lower grade influences the overall perception of the instructor. However, the idea of the article is to give an understanding that it is not a problem that needs solving but, on the contrary, it is an issue that has to be taken into consideration since teachers may also perceive this negative report as an opportunity to evaluate the effectiveness of the teaching methods. As an option, when giving a lower grade, the teacher might suggest a feedback or some other alternative methods of communicating the grade.

Another controversial aspect from the evaluation perspective was brought up by Willcoxson (2006), who researched the issue of grading teamwork. Since evaluating teamwork requires even more complex approach because it needs to consider not only individual but also group performance. The author highlights the importance of applying various grading methods including self-reflection, student's own evaluation of the teamwork and the work of the teammates. The observation of the flow of the teamwork is also an aspect that should be considered in the process of evaluation not only the final outcome of the work. So, the main aspect that Willcoxson (2006) wanted to emphasise is the complexity and difference of the teamwork's grading process comparing to just individual evaluation.

Duncan and Noonan (2007) explore the factors that have an influence on teaching grading practices including beliefs, views, attitudes and perception regarding grading practices in general and the ones that are used. Teachers reported that their assessment criteria are influenced by the opportunities for professional development, school regulations as well as district regulations. The study, where survey was conducted among 217 teachers and 17 follow-up interviews, revealed that teachers tend to include in their evaluation not only student's academic performance but also non-academic factors such as behaviour, effort and attitude. Consequently, Duncan and Noonan (2007) suggested that the inclusion of such characteristics can lead to the misunderstanding and misinterpretation of the grade. Additionally, when such factors are evaluated, the grading process results in inconsistency and may vary greatly from teacher to teacher and all of that, as a result, might have a negative impact on the student's learning outcomes.

The authors of the research suggest that the solution might be that the teachers are provided with the necessary opportunities for their professional development, where they will have a possibility to learn more about various assessment practices and will learn how to implement those practices in their assessment process so that it would become beneficial for the students' learning.

Similarly to Meuer (2006), Feldman (2007) argues in the study that student's rating of teacher's work might be a source of valuable information for the further improvement of teaching practices. Additionally, Feldman (2007) explores the issue also from the perspective of evaluation of teachers' effectiveness. The findings, nevertheless, are similar to the Meuer's (2006): when evaluating students are prone to include such subjective factors as instructor's personality, attitude, etc and thus such an evaluation might be only one of the tools of obtaining information on teacher's effectiveness and multiple other methods have to be used in order to obtain a more comprehensive picture.

Zoeckler (2007) explored in the study teacher's perceptions of the moral issues of the grading since grading is associated with fairness, honesty and equity. Even though the teachers believe that grading should reveal students' performance and academic achievement, they also agree that it is a subjective matter that usually influenced by teacher's own biases and beliefs. Those findings are similar to the ones suggested by Duncan and Noonan (2007) and the solution to this issue also consists in providing more professional trainings for teachers for them to be

more knowledgeable regarding various grading practices and in such a way it might be possible to develop more objective grading.

Standards-based grading might be also one of the solutions to the issue of subjectivity in the grading process stated by the above-mentioned authors. Scriffiny (2008) explores in the research this particular issue, stating that traditional grading practices are inconsistent and can not be used in order to understand students' level of knowledge. On the contrary, standards-based grading is way more straightforward and the evaluation criteria is clearer since it is bases on a particular learning goals and standards. Furthermore, such an approach to grading is used together with the feedback that also minimised misunderstandings of a particular grade. However, Scriffiny (2008) emphasises that even though this approach is more objective, it requires significant changes in the classroom and consequently might be challenging for teachers though not impossible when proper teacher trainings are provided.

Kulick and Wright (2008) conducted an experiment using Monte Carlo simulation in order to compare the difference in grade distribution in traditional grading and the curve method. Traditional grading is the one that uses fixed scale to evaluate students' performance while curve method is used to distribute grades in a particular classroom. The results were that when using the latter method, the students who usually receive lower grades are evaluated higher and visa versa with the students who receive higher grades. The authors came to the conclusion that curve method is not revealing the actual performance of students and thus more alternative methods such as standard-based grading are more effective and objective to use.

The research by Randall and Engelhard Jr (2009) demonstrated the usage of Rasch measurement theory in order to examine the consistency of the assigned grades. The aim of the study was to examine whether such factors as previous performance, ethnicity and gender influence the way teachers assign grades. The results were positive since teachers tended to assign lower grades to the students who received them previously. Moreover, lower grades were also assigned to the students from different minority groups even though their performance was on the high level. The authors suggest that the usage of Rasch measurement theory might be useful for identification of areas of concern in the evaluation process and might be used for the purposes of increasing validity and objectivity of grades.

Similarly to the previous studies, Bowers (2009) in his study emphasises the importance of having grade as one of many methods of evaluating students. Inclusion of peer assessment, self-assessment and performance can result in acquiring a broader vision of the student's overall achievement and will give the possibility to have a more objective view on an actual student's level of knowledge. The main idea that the author is emphasising is for the current teachers to rethink and reevaluate the importance and the role of the grades and the way they assess their students.

The study by Grimes (2010) examined closely the complexity of the grading issue by gathering information from middle school teachers through observation and the usage of descriptive statistics. The teachers reported to use multiple assessment methods however stated that there is lack of consistency of grading practices across different schools and even teachers. The concerns regarding subjectivity and biases during grading were also expressed by teachers themselves, Grimes (2009) concludes that more unified grading policies and practices need to be developed and taught to teachers and educators so that a more accurate grading that reflect student's actual achievement is used. This is the way that might lead to minimising teacher's subjective vision of the students' performance and knowledge.

The subjectivity issue was also raised by Swinton (2010) but was viewed under a different angle. The main idea of the research by Swinton (2010) is that effort should be a component of the the final grade. In this case, the grade is the summary of various factors. The drawback of this approach is that behaviour, attitude, effort and other factors are quite subjective and depend a lot on the teacher's personality. Thus, it represents not only the knowledge that the student acquired in the subject but also the teacher's subjective judgement. In this scenario, the grade can be misunderstood and misinterpreted by the students. Consequently, the students might even lose the motivation and interest in the subjective due to the lack of clarity of the grade. Moreover, students and teachers might consider grades as a source of teacher's power and significance and the grades are the main tool in such a situation. And there is no wonder that this attitude exists because for decades there was a big hierarchy between teachers and students and this hierarchy still exists in a lot of countries. So, it takes time and desire to change that attitude and transform it to a new form (Farias, G., Farias, C. & Fairfield, 2010).

Effort is obviously an important component of the learning process but simultaneously the evaluation of the effort may vary from teacher to teacher. Furthermore, the student's own evaluation of the effort might differ from the one given by teacher and when such a discrepancy repeatedly occurs, the student might lose motivation and interest towards the subject. So, the effort should be communicated between the teacher and the student but not in the form of a grade but possibly in a form of a feedback or comment.

2011 — 2022

Lekholm (2011) while investigating the effects of school characteristics such as size, teacher experience, and resources on students' grades discovered an interesting perspective. The study suggests that the grading practices that were used in a particular school had an impact on student's grades. Even though it was not a main objective of the researcher, the study concluded that there is a correlation between different grading practices and student's achievement and learning.

The research "The Case Against Grades" by Kohn (2011) is focused on the ways how traditional grading harmfully impacts learning process and emphasises the necessity of using more authentic forms of assessment. According to the study, traditional grading practices are too generalised: they are used to sort and rank students which can be misleading since such an approach does not consider students' individual learning style, circumstances and abilities. The inclusion of the above-mentioned characteristics is also questionable, since it adds more subjectivity. However, the point that the grading should not be too generalised is also a plausible argument. Grades are often used as a tool to motivate students but in many cases, it leads to the opposite effect because students start to learn a subject just for the sake of receiving a high grade, not for the sake of acquiring new and interesting knowledge. In such a way, mentioned Kohn (2011), students lose their intrinsic motivation, they become less creative and initiative and they show less interest in a subject in general. Based on that it can be concluded that when traditional grading practices applied the desire to learn for the sake of learning lessens and students are more driven by extrinsic motivation of receiving the highest grade.

The author suggests that the grades can be replaced by some alternative methods of assessment such as descriptive and specific feedback which can be combined with students'

own reflection on own work and learning. Meaningful feedback can give a solid ground for the students' improvement. The focus will be shifted from the final result, which is grade, to the process of learning. Such an approach promotes acquiring knowledge and teaches how to learn, which is an essential skill further in life as well. The usage of portfolios, self-assessments and peer assessments leads to the more engaged learning process, increased interaction between teachers and students and encouragement of students to take ownership of their learning, meaning that it is not only a teacher who decides to which extent student learnt a particular subject, but it is also a student who does that. Thus, a much broader image is acquired regarding the level of competence on a particular topic.

Marzano and Heflebower (2011) have been also researching the influence of traditional grading on the learning process. They argue that traditional grading practices tend to lack consistency and fail to evaluate student's performance accurately and thus can demotivate students. While outlining the main problems that are associated with the traditional grading, the authors suggest that more effective grading practices should be designed in order to improve learning process. The framework suggested by Marzano and Heflebower (2011) consists of providing accurate and meaningful feedback to both teachers and students, setting clear learning goals and provide opportunities for students' improvement. Often it is teachers who have a clear vision of how the course is structured and what is the flow of it while students have a general idea of that. However, knowing the structure can allow students to set much more specific goals for their learning process and it will facilitate the communication between teachers and students.

Specific individual feedback that provides actionable plan for improvement for students is another suggestion provided by the above-mentioned authors in the research "Grades that Show What Students Know". Furthermore, this kind of feedback is the most beneficial when provided throughout the whole term rather than just at the end of it. In such a way the students will have a chance to learn from their mistakes and an opportunity to improve their performance further. The article provides valuable examples of the schools that successfully implemented such a framework as well as gives guidance for the educators on how to use the above-mentioned practices in their educational institutions. The most important aspects that Marzano and Heflebower (2011) emphasise throughout their entire work is that the students' learning should be prioritised in the whole educational process. Using precise and consistent

grading practices supported by meaningful and descriptive feedback will result in students having the tools and motivation for their improvement.

Yesbeck's study (2011) explores grading from the perspective of which academic and non-academic factors are included in the grade. Despite teachers that participated in the study claim that non-academic factors such as attitude, participation, behaviour and attention, should not have an impact on the overall grade, they still tend to consider them when assigning a grade. One more factor that might influence the grade is the pressure from parents. So even though the teachers agree that inclusion of non-academic factors leads to the misleading grade since it does not purely represent an academic achievement, it is still hard to avoid considering them in their practices. The study emphasises the importance of the clearer assessment guidelines and policies as it is a way to ensure consistency in the overall grading practices.

Cox in the study "Putting Classroom Grading on the Table: A Reform in Progress" (2011) supports the idea that was also emphasised a lot in the previous studies: the importance of the feedback. The numerical grade shows very little of the information as for the skill mastery of a particular student, it just shows the number of points eliminating the consideration for learning and progress. Consequently, students' motivation and self-esteem can lower. The author's suggestion is that rather than judging student's mastery by the final result, it would be more beneficial for the learning to focus on the process. By demonstrating knowledge, receiving feedback and improving based on that feedback, the student has an opportunity to learn the subject and understand it. However, the implementation of such an approach to grading requires a collaborative effort from both school administration and teachers.

Many of the above-mentioned authors emphasise that grades might be treated as a source of teachers' power since they can be used as a reward or punishment and that is not the exact purpose of the grades. Brookhart (2011) in the research "Starting the Conversation about Grading" suggests re-examining existing grading practices in order for them to be such that reflect student's knowledge. The author mentions that even though grading practices have been examined for a long time they seem to still rank and sort the students based on their achievement. Brookhart argues that such practices as effort- or participation-based grading can seem harmless. Nevertheless, they can impact motivation and learning negatively. Approaching the grading from the perspective of what students know and can do, for instance

standards-based or mastery grading, could lead to the more accurate and meaningful grading that encourage students to take more ownership of their own learning. Moreover, that will have the tools for their own improvement. Similarly to Cox (2011), Brookhart emphasises the importance of communication between teachers, students, parents, administration and stakeholders in order to start implementing changes in the grading practices so that they can promote students' learning and growth.

Tierney, R. D., Simon, M., & Charland, J. (2011) conducted a study where 20 middle and high schools were involved. Those schools have implemented standards-based grading in their classrooms. The findings suggest that standard-based grading brings more fairness and accuracy to the assessment. Due to the fact that main principles of such an approach is clarity in setting the objectives, the usage of multiple assessment measures, focus on mastery of a particular topic, the learning is the main scope not the ranking and comparison with other students. However, teachers also pointed out some challenges that they faced with the standards-based grading. The need for more time, when developing and assessing learning objectives as well as balancing between being accurate and fair with the assessment while still being obliged to assign a grade. Moreover, there are still plenty of uncertainties when it comes to subjective matter of grading such as grading of an essay.

The authors of the article "Being fair: Teacher's interpretations of principles for standards-based grading" (2011) highlight the importance of further development of the standard-based grading since there are some promising results but there are still gaps that need to be filled with further researches and data. The collaboration and communication between teachers can be helpful and beneficial due to the fact that they can share their own insights and experiences.

Bower (2011) explored the multidimensional nature of the teacher-assigned grade. The main objective was to examine which factors teachers tend to include when they are assigning grades. Undoubtedly, teachers mention that they were assigning grades based on the students' mastery of a particular topic. Nevertheless, they acknowledged that such factors as effort and behaviour have also a significant impact on their judgement. This study showed that even with the clear intention of assessing only students' knowledge, it is challenging to remain purely objective and exclude effort and behaviour from the process.

Thorne and Cliffordson (2012) have also researched the factors that influence teacher-assigned grades but in addition to that the authors examined predictability of the grades on the criterion-based tests. The results revealed that teacher-assigned grades were moderately predictive possibly because of such non-academic factors as effort, behaviour and participation. Simultaneously, criterion-referenced tests were focused purely on students' academic mastery skills. The extent to which the grades were predictive depended on a subject. Thus, for English and Maths the grades were more predictive than for social sciences. The authors conclude that there is a need for using multiple assessment tools in order to receive a broader understanding of a students' level of knowledge. One more suggestion was to align teacher-assigned grades more with the learning outcomes in order to minimise the inclusion of non-academic factors.

The article by Knaack, Kreuz, and Zawlocki (2012) is focused on investigating standards-based grading and its effectiveness. For the purpose of data collection, a study was conducted in a high-school English classroom where standards-based grading has been implemented for one semester. Such an implementation aims to assess students' mastery of a specific skills and knowledge rather than the overall performance and in such a way students' weaknesses and strengths can be identified and improved if needed.

The results showed that the students viewed standards-based grading as a more informative way of assessment compared to the traditional grading. However, the authors mention that the implementation of the standards-based grading requires a significant amount of time and effort from the teachers. So, when considering using such an approach to grading the desire itself to improve grading is not enough. It is hard work that needs to be done in order to achieve results. According to the authors more research is needed to explore possible ways of facilitating the process of standards-based grading.

The author of the next study "A Better Grading System: Standards-Based, Student-Centered Assessment" (2013) also investigates the effect of the standards-based grading on students learning. Miller (2013) emphasises that traditional grading has a harmful impact on learning and motivation while standards-based grading supports learning process by the virtue of focusing on mastering the specific knowledge and skills. Similarly to Knaack, Kreuz, and Zawlocki (2012), Miller (2013) highlights that standards-based grading can help to identify which areas student is struggling with and can help to improve that area and gives an

opportunity to track the progress over time. The author acknowledges that the approach comes with some challenges when implementing. However, with effort and collaboration between teachers, administration and students it is not impossible to do, especially considering that standards-based grading promotes students' learning.

Deep learning and its correlation to grades was researched by Campbell and Cabrera in the article "Making the mark: Are grades and deep learning related?" (2014). The study was focused on exploring whether such a correlation exists and what factors influence that correlation. With the help of mixed-method design, researchers discovered that students tend to receive higher grades when they are engaged in the deep learning. However, it was noted that it is a complex relationship which is a combination of multiple factors. For instance, students noted that it made a difference to them when the teacher promoted and appreciated deep learning. Some other factors that influenced the engagement in the deep learning is interest and motivation in the subject. Teacher's grading practices have also impacted the level to which the students engage in deep learning. Presumably because the way the teacher grade shows the amount of space where students are allowed to make mistakes and learn from them without worrying about the grades. So, Campell and Cabrera highlight that the correlation between grades and deep learning exist but it is not so straightforward, but it is rather the combination of a number of factors. They suggest that it would be beneficial for teachers to promote various approaches to deep learning and set clear expectations and goals. Additionally, meaningful feedback will promote further learning of the subject and will give a path for students' improvement in the field.

While Campell and Cabrera (2014) article is focused on a more specific issue of grading, Schneider and Hutt (2014) provide an in depth look on the A—F grading scheme and its implication over the time in their study "Making the grade: A history of the A—F marking scheme". The authors mention that this grading system dates back to ten 20th century and now it has become an integral part of the educational process with the aim of ranking students without thinking of the consequences that arise from that. Numerous teachers realise that such a grading scheme brings more inequalities and shift focus from actual learning to striving for a higher grade. The authors emphasise that throughout the years grading has been used as a tool to reward those who fit all the traditional norms of the education and punish those who does not fit. Schneider and Hutt (2014) also examine standards-based grading and mention

that while this method of grading that focuses on a skill mastery is more equitable and informative from the viewpoint of learning, it is not an ideal solution since it also has its drawbacks. Overall, the authors highlight the fact that traditional grading is deeply attached to the educational systems nowadays and in order for it to change a tremendous effort and rethinking is required.

Similarly to Campell and Cabrera (2014), the research by Soderstrom and Bjork (2015) emphasise the importance prioritising of deep learning over immediate performance gains. The authors highlight that often such notions as learning and performance are used as synonyms while they are separate entities that rather complement each other. Learning is a process of acquiring knowledge and new information while performance is an ability to demonstrate that knowledge. Acknowledging that, according to Soderstrom and Bjork (2015), can assist teachers to use various teaching practices in order to enhance helping students with difficulties during their learning process and help them to achieve higher results at the end.

Standards-based grading has been highly discussed by various authors above as an alternative grading practice that can help to provide meaningful feedback for students and can assist in shifting the focus from grades to the actual learning process, to the mastery of a particular skill. Hany, Proctor, Wollenweber and Al-Bataineh (2016) conducted a study with 85 teachers in the midwestern part of the United States. Participants were asked to share their perceptions and opinions regarding standards-based grading and its impact on learning and teaching. The results show that teachers mostly have a positive attitude towards standards-based grading mentioning the benefits of providing feedback as it allows to have a much clearer understanding of students' knowledge and their progress. Standards-based grading facilitated for teachers the process of finding the areas for students' improvement by providing more accurate feedback. However, the teachers from this study mentioned the same challenges as the teachers from the study of Knaack, Kreuz, and Zawlocki (2012): standards-based grading is more time-consuming, the need for the assessment but with the goal of providing meaningful feedback is also an area that teachers found difficult to balance.

Similarly to the above-mentioned authors, Hany, Proctor, Wollenweber, and Al-Bataineh (2016) suggest that standards-based grading is highly beneficial in comparison with the traditional grading. Nevertheless, the success of its implication depends on how this approach is treated on a higher educational level i.e. more training of the standards-based grading

should be added to the educational teachers' programs so that the teachers are more aware of this approach and have more knowledge on the practical implications of this approach in the classroom. The authors conclude that standards-based grading has a potential to improve both teaching practices and also students' performance.

Lee, Carberry, Diefes-Dux, Atwood and Siniawski (2017) continued the exploration of the standards-based grading. The purpose of the study was similar to Hany, Proctor, Wollenweber, and Al-Bataineh (2016) - investigation of the perceptions of the implementation of standards-based grading. The authors used mixed-method: surveys that consisted of open-ended questions and Likert-scale questions allowed to have a more general view of the issue while interviews gave an in-depth view of the teacher's experiences. The results were similar to Knaack, Kreuz, and Zawlocki (2012) and Hany, Proctor, Wollenweber, and Al-Bataineh (2016): teachers have a positive attitude towards standards-based grading since it led to better communication with the students, it helped to identify areas that need improvement and allowed to provide a more accurate feedback. Nonetheless, teachers were concerned that the implementation required more time, designing of the new assessment techniques was a challenge and grading criteria was confusing for students as well. The overall conclusion of Lee, Carberry, Diefes-Dux, Atwood and Siniawski (2017) was similar to the earlier works mentioned before meaning that standards-based grading has a potential to be an effective grading approach, but it requires precise and careful planning as well as more professional training.

Another angle of the standards-based grading has been suggested by Peters, Kruse, Buckmiller, and Townsley (2017). The authors investigated students' attitudes towards this grading approach. The study involved 15 teachers and 61 students from two schools in the United States where standards-based grading has been implemented and survey, focus groups and interviews has been used as methods. The findings showed that there has been a resistance among students towards standards-based grading. As the students were used to the traditional grading, they viewed this new approach as unfair and as the one that is inaccurate in assessing their abilities and effort. Students were concerned of an impact of the standards-based grading on their scholarships, college admissions and other future prospects. However, the research also indicated possible ways of reducing students' resistance. Those included proper communication and support between teachers and students. Teachers who dedicated

time to ensure that students understand how system works, provided frequent feedback and provided opportunities for students' engagement and reflection, as a result had students who were more open to the new grading system. The authors concluded that the implementation of the standards-based grading requires effort and training not only for teachers but also for the students as both parties should have a proper understanding of how the new system works. Peters, Kruse, Buckmiller, and Townsley (2017) highlighted that in order to implement standards-based grading successfully, clear and honest communication should be present between educators, students, administration and policymakers.

In the study "Student perceptions of the fairness of grading procedures: a multilevel investigation of the role of the academic environment" by Burger (2017) the focus was on the students' perception of grading fairness. The analysis was made by the virtue of data from National Survey of Student Engagement collected in the United States and the responses included 219,585 students from 626 educational institutions. The results of the study showed that there were individual and contextual factors that influenced students' perception of the grading fairness. Individual factors included self-involvement and engagement in the course. Those students who showed higher level of this factor perceived grading as fairer. Contextual factors included prestige of the educational program, the size of the classroom and selectiveness of the institution. The author summarises that students' perceptions of the fairness is the complex issue that should be addressed accordingly. Burger (2017) suggests that in order to improve the fairness of the grading meaningful and constructive feedback together with the consistent and clear grading criteria should be provided. Simultaneously, the author highlights the importance of the educational environments as it also complements the overall perception of the grading.

Anderson in the article "A Critique of Grading: Policies, Practices, and Technical Matters" (2018) provides an overview of and analysis of the grading policies, practices and the impact of technical matters in education. While providing brief historical overview the author highlights the fact that grades have always been a significant tool for communication between teachers, students and parents. Anderson (2018) argues that grading that is presented in the form of ranking students is outdated. The grades that are based on students' mastery and understanding of the topic should be used instead. To support the opinion the author provides the reasons that make traditional grading even more harmful for the students' learning.

According to Anderson (2018), traditional grading has an issue with being biased and that might have a negative impact on the students' performance and also their self-esteem. As a result, student's motivation towards the subject lowers and the value of the grade decreases. Teachers should be aware of the bias issue and try to strive for the fairness and eliminating of the subjectivity in the grading process as much as possible. To achieve this more professional training should be provided for the current teachers and more information regarding standards-based grading as well as the appropriate ways of providing constructive feedback should be added to the teacher-training programmes.

The impact of the technology has been mentioned by Anderson as well. The author explored its role together with the transformation of the grading that comes with the implementation of technology. Anderson argues that technology helps with the efficiency of the grading, however it decreases teachers' autonomy. So, technology might be a useful tool for assessing students as it brings more standardisation to the process but only when used with other grading tools and methods.

The significant idea mentioned in the article is that teachers have to be aware of the risks that come with the grading and be knowledgeable of a various alternative grading techniques that can be beneficial to students' learning. To do so the improvements to the teacher-training programs have to be made.

Another article worth mentioning when discussing grading methods is the one by McBride and Kahle (2018). The primary focus of their article "Standards-Based Grading: An Examination of Assessment Options" is the implementation of the standards-based grading in the classroom. Similarly to the previous study by Anderson (2018), this study criticises traditional grading due to its subjectivity and inaccuracy in providing a clear image of student's knowledge. The authors suggest that standards-based grading is a way to make grading more accurate and fair since it is focused on the mastery of a specific learning standards. The results of the study conducted by McBride and Kahle showed that on the one hand, teachers had a positive attitude towards the effectiveness of the standards-based grading. On the other hand, teachers are faced with the number of challenges such as alignments of the assessments with the standards, determination of the value of the particular standard as well as communicating the system to the students and parents who are used to the traditional grading practices. The authors highlight the idea of effectiveness of the standards-

based grading, benefits of using multiple assessment methods. Simultaneously, they realise that this grading method has its drawbacks and at first it might be demanding for the teachers to implement such a method in the classroom.

The article "Students tell us what good written feedback looks like" by Voelkel, Varga-Atkins, and Mello (2020) gives an in-depth view on the students' perceptions of the written feedback. Since feedback is considered to be one of the keys to the successful grading, this study gives suggestions for what kind of feedback is appreciated by students in the way so it can complement and improve learning outcomes. Students value feedback that is timely, precise and has suggestions for further improvements. They highlight that personalised feedback with an emphasis on the work that they receive the feedback for as well as the mistakes that they have made there is the one that helps them in their further learning as they become aware of the areas that they need to improve. The authors suggest teachers to consider such a perception so that their feedback is of the most value to the students. Moreover, Voelkel, Varga-Atkins, and Mello (2020) suggest that the teachers ask students regularly to reflect on the feedback that they have received and the ways it can be improved from the viewpoint of students. Overall, the article highlights the importance of giving feedback even though it might be time-consuming as it can result in students' increased motivation and better learning outcomes.

Ismail and Tini in their article "The effect of formative assessment on student's grades in learning English" (2020) investigated the correlation between formative assessment and graded by the virtue of experimental group of 40 students and control group of 40 students as well. During the study both groups have been taught with the same materials and methods, but experimental group had been given formative assessments in the form of self-assessments and feedback while control group received summative assessment in the form of test.

The results reveal that experimental group has received higher grades and students from this group showed better learning. Formative assessment gave those students tools for the improvement of their knowledge and grades. Moreover, students have perceived formative assessment positively as it helped them to identify their strong and weak areas, and self-assessment has been helpful in setting goals and tracking their own learning progress.

Ismail and Tini (2020) emphasise the positive outcomes of their study and suggest that formative assessment might be an effective tool to improve grading and make it more

informative. Additionally, formative assessment encourages students to participate actively in their own learning which has been shown to be beneficial for the overall learning outcomes.

Lipnevich, A. A., Guskey, T. R., Murano, D. M., & Smith, J. K. (2020) investigated the variation in the grading criteria among different instructors and courses from 23 educational institutions in the United States. They found a significant variation in grading criteria across institutions, departments, courses and instructors. Grading criteria was influenced by a variety of factors such as student's performance, mastery of a subject, effort, improvements as well as institutional culture and teacher's personal philosophy. Grading criteria also varied based on the type of the assignment that was graded: whether it is an individual assignment or a group work. Even though it is natural to have different grading criteria for different types of assignment, the authors suggest that a greater consistency is needed in the educational institutions. Providing clear and consistent guidelines for the grading criteria is a way to make assessment more comprehensive for both teachers and students. Moreover, it might be beneficial for the educational institutions to provide teachers with more professional developmental opportunities so that they can improve their grading methodologies.

The article "Going 'Grade-Free'? – Teachers' and Students' Perceived Value and Grading Preferences for Holistic Competency Assessment" by Chan and Luk (2022) investigates how the shift in assessment practices that goes beyond traditional grading has a potential to better capture students' competences and prepare better for the future workplace. The authors discuss the limitations of the traditional grading which is primarily focused on the content mastery rather than skills and competencies and that can negatively impact the motivation towards learning in general. The authors suggest that the more holistic approach towards grading that includes multiple forms of assessment such as portfolios and self-reflection can provide a better understanding of student's abilities and knowledge. The study by Chan and Luk (2022) used survey, where 57 teachers and 518 students were participants, as a method for data collection. The results show the positive attitude among both teachers and students towards a more holistic approach to grading where grading and non/grading assignments are combined, and the feedback is provided for the purpose of further improvement. Authors suggest that the implication of a more holistic approach has a potential to result in improved motivation which is beneficial for the overall learning outcomes.

Zairjanovna, Mamadjanovna, Salaxiddinovna and Tulaboevna (2022) in their article also argue about the negative effects of the traditional grading but from a psychological viewpoint. They mention that traditional grading impacts negatively students' motivation towards a particular subject and learning in general as well as brings more anxiety and fear of failure. Zairjanovna et al. suggest that the grade-free system with the focus on feedback creates a more positive educational environment which promotes students' motivation and reduces stress of failure. The authors conducted an experiment where 120 students in Uzbekistan were randomly divided into experimental and control group and received feedback-based grading and traditional grading respectively.

The results showed that the experimental group had higher levels of engagement, motivation and academic achievement compared to the group with the traditional grading. Grades-free environment allowed students to focus more on their learnings and their progress based on the personalised feedback rather than fear of receiving lower grade.

Zairjanovna et al. highlight the positive outcomes of the grades-free learning system and emphasises its value as it promotes student's motivation and academic achievement. They also discuss the implications of the findings in education but also note the need for further investigation of the topic and various assessment techniques and their impact on the learning outcomes.

Overview of the main grading themes

Historical overview part that was presented in the section above intended to show the development of the ideas and views towards grading. Some of the themes have been reoccurring throughout the time. The intention for this part is to provide a schematic overview of the most prominent topics in order to demonstrate their chronology and timespan. It is worth mentioning that the development as well as the timespan were not linear. That is why the figure below will show the density of the research in the particular time period. The main intention was to show the major themes that has been present in the researches of a particular time period. The more detailed table will be presented further in this section. The categorisation, demonstrated in both tables, is going to be used as one of the frameworks for the methodological part where results of the practical work will be communicated.

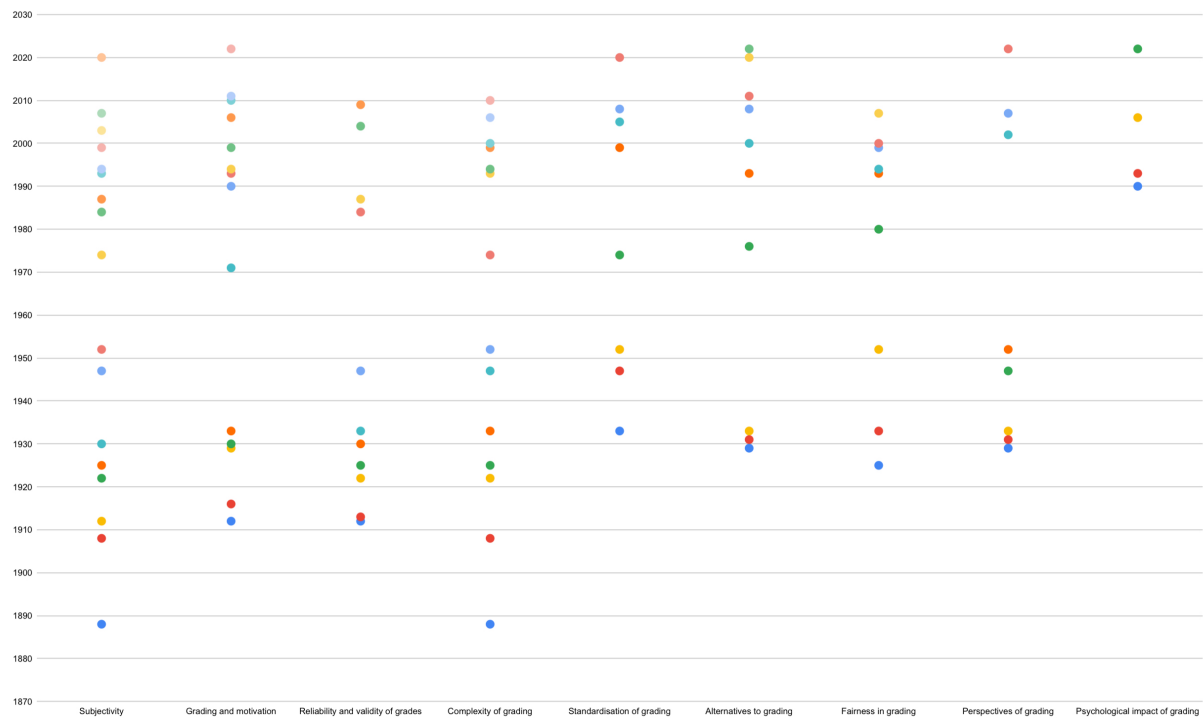


Figure 2. Graphical representation of the density of the major grading themes.

Figure 2 demonstrates that given the time period from 1800 until 2022 there have been ten main topics that have been researched. Such topics as subjectivity in grading, grades and student's motivation, reliability and validity of grades, complexity of grading, standardisation of grading are the ones that has been researched the most and the questions regarding them remain open until nowadays. The first discussion of subjectivity takes its roots in the work of Edgeworth (1888). Complexity of grading as well as reliability and validity and the need for the standardisation has started in the beginning of 20th century with the works of Edgeworth (1908) and Meyer (1908). In 1921 researches in the educational field started discussions regarding various alternatives to grading. Yet nowadays education continues searching for those alternatives that will be beneficial for the learning process. Fairness in grading, perspectives of grading, psychological impact of grading and challenges in implementation of different grading practices — all those topics started their way in the middle of the 20th century or even earlier. For example, Edgeworth in his research in 1888 mentioned the influence of stress on grading, emphasising that if student is under the stress during the exam, it will have an influence on the final grade. Considering that, the grade demonstrates not purely student's knowledge, but it becomes a combination of other factors in it. Thus,

psychological aspect of grading has its roots deep in the history of grading, but the most of research still has been made in the 21st century.

The Table 1 below depicts the more precise image of the occurrence of a particular grading themes throughout the history. Moreover, the table proves that all of the topics appeared in the earlier works and then were changed by some other topics, and then they appeared again possibly with another perspective.

Table 1. Representation of the occurrence of the main grading themes in the researches

| № | Themes | Researches |
|----|------------------------------------|--|
| 1. | Subjectivity in grading | Edgeworth (1888), Meyer (1908), Starch and Elliot (1912), Silberstein (1922), Hulten (1925), Eells (1930), Swineford (1947), Carter (1952), Collins and Nickel (1974), Marsh (1984, 1987), Brookhart (1993, 1994, 1999), Smith (2003), Duncan and Noonan (2007), Zoeckler (2007), Lipnevich, A. A., Guskey, T. R., Murano, D. M., & Smith, J. K.(2020) |
| 2. | Grades and student's motivation | Starch and Elliot (1912), Crooks (1933), Morrison (1926), Odell (1929), Crew (1930), Hillbrand (1930), Holmes (1971), Janzow and Eison (1990), Brookhart (1993, 1994, 1999, 2011), Meuer (2006), Farias et al.(2010), Kohn (2011), Marzano and Heflebower (2011), Cox (2011), Zairjanovna et al. (2022) |
| 3. | Reliability and validity of grades | Starch (1913), Starch and Elliot (1912), Silberstein (1922), Hulten (1925), Eells (1930), Crooks (1933), Swineford (1947), Marsh (1984, 1987), Woodruff and Ziomek (2004), Randall and Engelhard Jr. (2009) |
| 4. | Complexity of grading | Edgeworth (1888), Meyer (1908), Silberstein (1922), Hulten (1925), Crooks (1933), Swineford (1947), Carter (1952), Ebel (1974), Brookhart (1993, 1994, 1999), McMillan and Nash (2000), Willcoxson (2006), Grimes (2010) |
| 5. | Standardisation of grading | Crooks (1933), Swineford (1947), Carter (1952), Collins and Nickel (1974), Brookhart (1999), Allen (2005), Scriffiny (2008), Lipnevich, A. A., Guskey, T. R., Murano, D. M., & Smith, J. K.(2020) |
| 6. | Alternatives to grading | Crooks (1933), Lund (1929), Madsen (1931), Simon and Bellanca (1976), Winter (1993), Guskey (2000), Scriffiny (2008), Kulick and Wright (2008), Kohn (2011), Marzano and Heflebower (2011), Ismail and Tini (2020), Chan and Luk (2022) |
| 7. | Fairness in grading | Hulten (1925), Crooks (1933), Carter (1952), Abrami, Dickens, Perry, and Leventhal (1980), Brookhart (1993, 1994, 1999), McMillan and Nash (2000), Zoeckler (2007), Duncan and Noonan (2007) |
| 8. | Perspectives of grading | Crooks (1933), Lund (1929), Madsen (1931), Swineford (1947), Carter (1952), Guskey (2002), Duncan and Noonan (2007), Chan and Luk (2022) |

| № | Themes | Researches |
|-----|---------------------------------|--|
| 9. | Psychological impact of grading | Janzow and Eison (1990), Winter (1993), Meuer (2006), Zairjanovna et al. (2022) |
| 10. | Challenges in implementation | Crooks (1933), Cizek, Fitzgerald, & Rachor (1995), Brookhart (1993, 1994, 1999), Scriffiny (2008), Bowers (2009) |

It is noteworthy that practically all of the topics have been presented to some extent even in the end of 19th century — beginning of 20th. So, basically all of the above-mentioned topics have their roots in 19th century or probably even earlier, but there has not been found evidence of that due to the lack of digitalised records of the researches. So, the above-mentioned categorisation was made in such a way that it shows the development of different grading ideas. For instance, in the end of 19th century most of the researches were focused more on the problem of subjectivity of grading, however, it does not mean that other topics have not been discussed. While getting acquainted with the earliest works it is possible to notice some mentions regarding all of the topics, which makes them multi-dimensional. That also proves that grading has been an issue for a long time. And even though new approaches arise, those very first problems remain unresolved.

3. Methodology

This part is intended to explain the main methodology that have been implemented in this thesis work as well as describe the results of the practical part of the work. This part will give an overview of the objectives of the thesis work as well as state the hypothesis that were formulated in a way to correlate with the objectives. Furthermore, the data collection instrument, its design and the choice of the participants will be mentioned and justified.

Objectives and Hypotheses

The main objective of this thesis work is to discover possible correlations between grading and student's learning, the ways they influence each other in both positive and negative way. In order to do so students' opinions towards this will be collected.

The following hypotheses and sub-hypothesis were formulated in a way that helps to research the objectives from the multiple angles for the purpose of gaining the broader overview of the grading issue.

Hypothesis 1: Students experience uncertainties regarding the interpretation of the grades due to the lack of understanding of the components that are included in the grade.

Hypothesis 2: There is a variation between and across students in perception of the grading.

Hypothesis 3: Students tend to perceive feedback together with the grade as an effective method in improving their grading experience and the overall learning outcomes.

The importance of observing whether the ***hypothesis 1*** is supported or not lies in the fact that if the students are confused and unsure what the grade represents and which components are included in the evaluation process, it means that it might be hard to use grades as a valid tool in the learning process or to estimate the extent to which grades illustrate student's actual performance and knowledge. To check the hypothesis the answers from Likert scale questions and open-ended questions will be analysed and presented further in the results section.

Hypothesis 2 aims to explore the reactions of students towards different situation scenarios. Exploring whether this hypothesis is supported or not will allow to observe possible discrepancies across the persons, meaning that the student might consider fairness of the grade that does not meet their expectations differently depending of whether it was higher or lower than was expected. In the results part 2 situations where the students are asked to describe their attitude towards the grade that is lower and the grade which is higher then they

expected will be presented. The purpose for presenting those situations is to show discrepancies in perception of fairness of grading among students and highlight the complexity of the grading issue and the importance of its further research. Additionally, the variation of perception of the evaluation between students will be analysed with the aim of demonstrating the complexity of the grading that goes both on the personal level and a broader one.

Hypothesis 3 aims to collect students' experiences regarding their attitude towards feedback that is provided together with the feedback. It could reveal the way to improve the grading experience together with the learning as such. Feedback has been stated by the numerous researchers throughout the time as a beneficial instrument for the improvement of the learning process and a way to avoid misunderstanding between students and teachers. For instance, Siberstein (1922) and Crooks (1933) were mentioning that providing clarification for the grades results in minimising misunderstandings between teachers and students. Carter (1952) and Abrami et al. (1980) emphasised that feedback together with the grade can improve student's overall grading experience. Simon and Bellanca (1976) while exploring various alternative grading practices concluded that feedback has a positive effect on student's performance and understanding of the learning material. Further researches, such as Jansow and Eison (1990), Winter (1993), Brookhart (1999), Guskey (2000), were also exploring different grading techniques but the method that all of them found beneficial for learning and the one that has a potential to reduce stress during evaluation is feedback. There is plenty of researches which are in favour of having a feedback together with the grade but one unique aspect that will be analysed and further highlighted based on the students' responses is the way how the feedback is perceived and described as feedback is also a form of grading and it can also be perceived both positively and negatively. There were no direct questions regarding that but by the virtue of multiple open-ended questions I will probe to find answers to that issue.

Data collection instrument

Questionnaire (that can be found in Appendix 1) has been used as a method for data collection for this work aiming to support or contradict the hypotheses in this master's thesis as well as answer the research questions. This method has been used due to the variety of reasons. First

of all, it is scalability—it was possible to distribute the questionnaire to a wide variety of participants and receive more data than with any other method. Secondly, due to the fact that the questionnaire was anonymous—no names or other personal details were required in the questionnaire—the respondents were able to freely share their thoughts regarding their grading experience and their feelings towards that. All the above-mentioned combined enabled to gather quantitative data, which will be analysed statistically and the broader picture of the perceptions and experiences will be presented. Furthermore, with the questionnaire as a method for data collection it was possible to have different types of questions such as Likert scale questions, open-ended questions, multiple-choice questions and that helped to explore the same issue but from a different angle.

The questionnaire was designed in a way to gather the information about perceptions and experience with grading from different perspectives. The questions were formulated in a way to cover those aspects of the research which are focused on the perception of fairness of the grade by students, the ways to avoid misunderstandings between teachers and students, whether grading is an essential part of the learning progress and whether grades or learning considered by students to be of an importance. The effectiveness of the design of the questionnaire has been checked according to Dilman, Smyth and Christian (2014), who mentioned key elements of the questionnaire in their book “Internet, phone, mail and mixed-mode surveys: the tailored design method.” According to Dilman et. al. (2014) the elements of the effective questionnaire are the following:

1. The usage of the clear language and clear instructions: judging from the responses received, which were precise and directly answered, it can be concluded that the questions and instructions were understandable.
2. Combination of closed- and open-ended questions: the questionnaire was designed in a way to have a balance between those two types of questions in order to facilitate answering for respondents and to have an opportunity of receiving both quantitative and qualitative data.
3. Logical questions order: to research multiple aspects of grades and grading, the questionnaire is divided into the series of questions of a similar topic and include direct and indirect questions regarding similar issues for the purpose of obtaining broader vision of students’ perspectives as well as noticing discrepancies in the answers. The questions

were divided into three categories: a) student's understanding of the grades—this category is aimed at discovering whether the first hypothesis is supported and whether the students have an understanding regarding what the grade shows; b) students' perceptions of the grade—this set of questions is focused on gathering information regarding students experiences towards grading and fairness in the assessment process; c) correlation between grading and learning—the purpose of these questions is to discover the ways students connect grading and learning, whether they consider grading to be an integral part of the evaluation process. Those three groups follow one another and complement each other. In the beginning the questions are more general, which is aimed to facilitate the process of emersion into topic for respondents. Closer to the end there are more open-ended questions which require more reflection on the topic.

It is also worth mentioning that the variety of different types of questions in the questionnaire allows researching the same issue from the different angles. For instance, the Likert scale questions included direct statements where participants needed to express to which degree they agree or disagree to the statement regarding fairness of the grading. Further, the situations have been presented regarding the same issue but participants were asked to elaborate on those. Such an approach allowed to receive better and broader understanding of the way students perceive grades and grading. In the open-ended questions respondents are able to share their thoughts, situations that they experienced etc. As a result, during the analysis of those responses it will be possible to observe possible correlations between similar questions.

Considering the above-mentioned principles stated by Dilman et.al (2014) and how the questionnaire used as the method for data collection in this thesis align with them, it is possible to mention that it was designed effectively and thus can be viewed as a valid method for collecting data.

Overall, the questionnaire's main intention is to gather the complex and multi-dimensional information regarding students' perceptions and experience of the grading as well as discover their suggestions on how the grading might be improved.

Participants

International students at the University of Turku have been used as participants for this research. The questionnaire has been answered by 42 international students of the University of Turku and the results that are presented below will be based on those answers. Their age was of no importance but it was crucial to obtain data particularly from the international students as their grading experience is versatile and they experienced various educational systems. This gives the opportunity to have a broader overview of the grading issues and highlights the most prominent ones. If the issue of subjectivity, for example, emerges in the answers of different respondents, that means that this issue is widespread and occurs in different countries and educational systems. It will also align with the researches that have been made throughout the time, starting from Edgeworth (1888) and the fact that he admitted the inaccuracy of grading and its variability from teacher to teacher. The more recent studies such as Brookhart (2016) and Lipnevich et. al (2020) also emphasise the issue of subjectivity in the grading process. So if all those sources combined: responses to the questionnaire and theoretical background from different time periods, visualise the same issue, then it would make it possible to highlight that issue even more and emphasise the importance of its further investigation.

4. Results

As it was mentioned in the section above, questionnaire has been used as a method for data collection. Below, some of the main findings will be presented. Firstly, the results will be analysed in a way to observe their possible alignment with the hypothesis and later whether they answered the research questions of this work.

Moreover, the literature overview part allowed to categorise the researches by the most prominent themes and the results will be categorised accordingly in order to observe whether the themes remain relevant nowadays. Additionally, those themes will be overviewed through the prism of their beneficial and detrimental aspects based on findings from both literature overview part and this study.

This part will be divided in the same sections as were used in the questionnaire: understanding of the grade, students' perceptions of the evaluation and the interconnection between grading and learning. The hypotheses, that were stated in the methodology section, and the research questions, that were mentioned in the introductory section but will be presented once again together with the results, can be divided in a way that they fit into those categories. This is done for the purpose of structuring gathered information and to connect all of the parts. Moreover, there will be additional subsections for the purpose of highlighting the crucial aspects of grading mentioned by the participants of this study, as well as for gaining the broader image of the grading situation.

Students' understandings of the grade

The term "understanding" in this case can be defined as an awareness of which aspects are included in the grade. This subsection aims to check whether the *hypothesis 1*, which is "*Students experience uncertainties regarding the interpretation of the grades due to the lack of understanding of the components that are included in the grade*" is supported or not based on the analysis of the collected data. Researching this aspect is of importance because in order for the grade to be a valid measure of one's knowledge, and 44% of the participants agree or totally agree with that, there should be an overall understanding of which parameters are included in the grade and what the grade represents.

When asked directly whether students are aware of the reasons for receiving a particular grade, majority of the participants, to be more specific 83%, either always or usually know

those reasons. Another question asked whether the teacher explains criteria for the course and the Figure 3 below demonstrates the distributions of the answers. According to that 81% of teachers always or usually clarify the parameters which will be considered during evaluation.

Teacher explains the evaluation criteria for the course
Number of respondents: 42

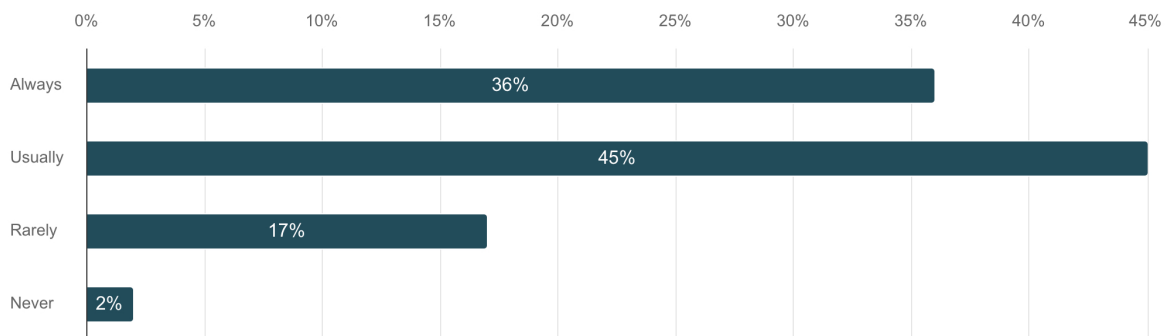


Fig.3. Responses to whether the teacher explains the criteria for the course.

Additionally, 86% of the students responded that they can ask for teacher's clarification in case they are unsure why a particular grade has been assigned. However, 14% of the students who would not ask the teacher for clarifications for the grade and as a justification for their answer, replied that they either thought that asking for clarification from the teacher is disrespectful or impolite, some did not know that they were allowed to ask anything like that from the teacher, or were feeling shame for asking. Even though it was a minority of participants who mentioned those aspects, it is still worth mentioning because it demonstrates that the understanding of the grade might be influenced also by cultural differences and what is normal in one culture might be unacceptable in another.

Based on the information above, it seems that the majority of students do not experience uncertainties and have an understanding regarding the components which are included in the grade. So, hypothesis 1 is not supported.

However, during the analysis of the open questions and more specifically the one that required to describe which changes should be made in order to improve grading, majority of respondents mention communicating grading and course criteria clearly which leads to the conclusion that even though teachers state those, students still feel that they are not completely satisfied with the way those are delivered. Based on that, it is possible to state that the *hypothesis 1 is partially supported* since students obviously have some kind of

understanding of what the parameters that the teacher takes into consideration are but it would be beneficial to know it more clearly.

Students' perceptions of grades and evaluation

The previous section revealed that students have an understanding of the grades to some extent and this section aims to study the perceptions of the grades and the overall evaluation process. While researching this, *hypothesis 2* which says that “*There is a variation across and within the students in perception of the grading*” will be checked as well as two research questions: on the ways of minimising discrepancies in perception of grading between teachers and students, and whether the perception of fairness is influenced by the grade being lower or higher than they expected, will be tried to be answered based on the collected data.

Firstly, the idea of fairness from the students' perspectives will be analysed. In the distributed questionnaire, the respondents were asked two questions concerning their feeling of fairness when receiving a lower or higher grade than expected. Similar issue has been investigated by Holmes (1971), but with a slightly different angle. In the study the level of student's involvement has been researched and it was found that the students who expected higher grade were showing higher level of involvement. Simultaneously, no strong correlation has been found between lower than expected grade and involvement.

This study, instead of the involvement, was focused on the feeling of fairness which the students associate with lower and higher than expected grade. Generally, both of those situations should feel unfair because the expectations and the reality did not match. However, the results showed a rather different image: 81% of respondents answered that they were in a situation when they received **grade lower** than they expected, of which 74% felt that this was unfair. 62% of respondents were in a situation when they received **grade higher** than they expected, out of which only 23% marked that this was unfair. So, based on these numbers people tend to consider grading fair or not based on what the consequences of those grades are: receiving lower grade affects negatively the picture of the overall performance while the higher grade has a positive effect. The hypothesis 1.1 was aimed to check whether the students have a coherent understanding of the grade themselves. And these quantitative results showed some discrepancies with that.

To explore this aspect even further, the respondents were also asked to elaborate on two situations where expectations regarding the grade and the received grade did not match. Further, two tables with some of the responses will be provided together with some explanations. Their main aim is to support or contradict the quantitative results on the issue of fairness when the grade is higher or lower than received.

Table 2, which is presented below, demonstrates the responses of the students to the situation where the grade that they receive is lower than they expected. Some of the common expressions that were used by the respondents included losing interest in the subject and even studies in general, losing motivation, feeling bad, disappointed, anxious. Similar findings have been made by Swinton (2010) whose study showed that when students receive the grade that is lower than they expected, they might feel demotivated and even experience cognitive dissonance which then results in the negative psychological impact of the grading in general. Some also mentioned that one of the reasons for such a situation might be that the clear criteria were not presented by the teacher or that those guidelines were not fully understood by the students. However, in the previous section, participants stated that in the majority of cases, 81% to be more specific, teachers clearly communicate the criteria for evaluation. That shows a discrepancy which supplements the support of the hypothesis 1 and demonstrated how students actually feel regarding it.

Table 2. Discrepancy between expected higher grade and received lower grade: students' responses.

| You wrote an essay that you really liked. You used good, relevant literature resources and put a lot of effort into it. You are satisfied with the result of your work and expect to receive a good or excellent grade. Finally, the result comes and you receive a satisfactory grade. How would that make you feel? What would be your first thoughts? Would you ask a teacher for a reason for that grade? Why/why not? |
|---|
| I would think that perhaps there were better answers than mine, and that maybe I did not fully answer the question. I would maybe do a bit of literature search regarding the topic and if I would find no reason to warrant a lower grade, I would contact the teacher who graded the essay and try to discuss and reflect on what I could have done better |
| I would try to re-read the assessment criteria again and try to find the gaps between the criteria and my work. I have never asked for a reason or justification for my grade, as I am not and I do not want someone to think that I am grade oriented student. |
| Disappointment is the first feeling I have. Depending on how approachable I feel the teacher is, I might ask the teacher for some feedback. |
| I would be mortified and consider giving up on my studies if trying my best only reached a satisfactory grade. I would not ask about reasons as I would be so ashamed. |
| I would feel very disappointed and think how could this be - feelings of disbelief in the very beginning, wondering what is wrong with what I thought initially. There may be feelings of giving up, eg. what's the use of putting so much hard work into it. Or I may rationalise in my mind that a satisfactory grade is also good enough, what's important is I think I have learnt what I need to learn and I'm happy with my learning process. Yes, I would ask the teacher for a reason to find out what's wrong. |
| At first, I would be wondering what has gone wrong. I would feel sad that I have put so much effort in and it did not "pay" back sufficiently. In such a case, I think I would ask the teacher for further explanation of the grade to understand what has gone wrong. In this way, I could also learn for my future writings. |

Noteworthy, that not that many respondents would ask for the clarification for the grade: some were feeling ashamed, some mentioned that the teachers are not reachable. Considering that and comparing with the results from the previous section, it is noticeable that there is a slight discrepancy: in theory the students are willing to ask teachers for clarification of the particular grade but in practice, there is an uncertainty regarding the way teachers will perceive this kind of question. So, based on the answers, students would stay dissatisfied and disappointed with their grades that did not match their effort and later their expectations but at the same time they would not do anything to use that grade as a base for their further learning. Another table, table 3, will demonstrate the situation regarding grading and some responses to it. Based on that table and on the other responses that were not included in it, if the situation is reversed meaning that the expectation was lower than the received grade, the students mostly feel happy, wonderful and pleasantly surprised and there is a very few responses where students mention that they will ask for the clarification of the grade from the teacher. Some students even consider higher grade as a gift: "...don't look a gift horse in the mouth...". In such a situation the whole idea, where grade is the representation of student's knowledge,

Table 3. Discrepancy between expected lower grade and received higher grade: student's responses.

| You wrote an essay that is OK. You did a decent job but you do not expect to receive a high grade. Finally, the result comes and you receive an excellent grade. How would that make you feel? What would be your first thoughts? Would you ask a teacher for a reason for that grade? Why/why not? |
|---|
| It would be good surprise for me but I would not start looking for a reason. Most superficially I would think that the rest of the essays were really poor in quality. |
| Wonderful. I would not ask the teacher for a reason because first you don't look a gift horse in the mouth and second because as long as the teacher finds it excellent I'm OK with it. |
| Great, I would maybe ask but probably would not bother if I already got a good grade. |
| I would not ask about anything because it is against my interests. Most likely, I will think that this is an example of (unfortunately) different demands in different courses. Students know that some courses are easy to pass, while some others are more difficult, although the amount of ECTS is the same. So, the same applies to essays, in practice. |
| Of course quite surprised, but also satisfied that besides of probably some difficulties, I understood what the teacher wanted to transmit and learn from it. |
| I would think that I again underestimated myself, which happens from time to time. I would not ask a reason for the grade as it does not matter for me. |

becomes almost meaningless because it is impossible to predict it based on a certain factors. One more aspect worth considering is that some respondents mention that there is a probability that they underestimated their work, which undoubtedly could be the case, and in order to be sure about that, the student still needs some feedback from the teacher. In such a

way there will be a ground for further improvement. The open-ended questions showed the vivid difference in perception of the higher and lower than expected grade. Both statistical data and the students' perceptions allow to answer one of the research questions regarding whether the grade is unfair to students only when it is lower than they expected.

Responses to these two situation make it possible to answer the research question regarding the discrepancies in students' perceptions of fairness when being over or under graded. So, students do perceive the concept of fairness differently based on whether the grade is higher or lower than they expected meaning that it is not only about fairness but also about what is beneficial for the students.

To explore further students' perceptions of fairness, the respondents have also been asked to describe fair grading from their perspective. When describing fair grading most respondents answered that they view grading as such where the clear criteria and requirements are provided. They also mentioned the importance of being objective, unbiased and transparent during the evaluation process. There were multiple answers pointing to the need to make the grade not only about the final work but about the performance during the whole course as well as effort. However, in such a situation, the objectivity might be corrupted.

The above-mentioned aspects have also been mentioned by participants when describing the ways to improve grading and minimise misunderstanding between teachers and students throughout the evaluation process. From the respondents' answers it became possible to conclude that there are several components that students view as crucial in the evaluation process:

1. Transparency and objectivity in the process of grading.
2. Feedback that goes either together with the grade or after a certain assignment that is ungraded.
3. Clear instructions and structured guidelines for the course and for the assignment.
4. Openness for communication: possibility to ask for clarification without fear of receiving lower grade or possibility to ask for help or suggestion.

All those aspects are often already a part of the grading process. However, based on the fact that lots of respondents had those aspects mentioned, it means that there is still a space for improvement and for the better communication and applying those principles to the evaluation process might improve the communication and understanding between teachers

method in improving their grading experience and the overall learning outcomes” is supported or not. Additionally, the answers to such research questions as whether the students prioritise grades over learning and consider them to be a valid measure of their knowledge and whether the grades are considered to be an integral part of the learning process will be attempted to be found.

Firstly, a noteworthy discovery has been made as for whether students consider learning to be more important than grading. The interconnection between grading and learning and how they influence each other has been the major focus of this work and thus a lot of questions in the questionnaire intended to observe this relation both through direct and indirect questions. In the questionnaire there were a few questions which were meant to observe whether the students understand and distinguish the difference between learning and grading. The analysis is done respondent by respondent as it allows to notice possible discrepancies in the answers within one person.

Question 20 asked the students to agree or disagree with the presented statements. Statements *“When taking a course, I focus on my learning and my improvement rather than on the grades”* and *“Receiving the highest grade is the best motivation for me”* are now being analysed. Presumably, if the person’s main focus is learning and improvement rather than grades then receiving the highest grade should not present itself as a best motivator to a student. However, 15 respondents, which is approximately 36% of a total number of respondents, answered *Agree or Totally Agree* to both of these questions. It is less than a half of the respondents yet it is still a significant amount of students.

The Figure 5 below shows visualisation of the two questions discussed above and depicts a few more statements worth of attention.

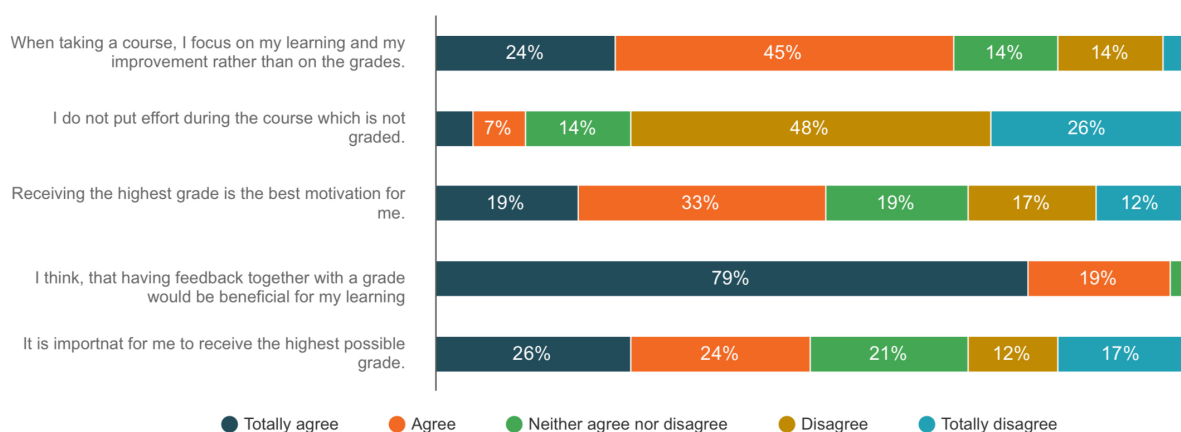


Fig.5. Graphical representation of the responses regarding interconnection between grading and learning

All those statements aim to discover the connection between grading and learning. The statement “*I do not put effort during the course which is not graded*” was the question which intended to observe the possible prioritisation the grades over learning. Judging by this questions students tend to put effort into the courses which are not graded. 74% of respondents disagreed with the above mentioned statement. The Figure 5 demonstrates the graphical representation of the above mentioned statement and the distribution of the answers. However, the further question which asked to describe how the students would feel if there was no grading revealed a different perspective and Table 4 below shows several examples. Overall, students state that they would lose their motivation or will not put effort as every student is the same when the course is just pass or fail and these answers contradict the ones that were presented in the Figure 5.

Table 4. Responses to the question regarding absence of grading

| |
|--|
| Only pass or fail? Great, I can do the bare minimum and pass a course, free credits. On the other hand, without grades, and everyone passing, we are saying that all of our knowledge are the same and we are in the same level in the learning process and we have the same abilities, and skills that are required to be professionals in our field. Is that true though? No, so I would also feel unfair if I put great effort in a pass but someone else does not and they just pass, that's why I wouldn't also place much effort in my studies then. |
| Interesting question, I think I would be less motivated. I love learning but getting essays written and doing exams require struggle and intention and if there was not the reward of a good grade I am nit sure I would get things done and I am sure I would not put as much effort in them. |
| We have been given grades so far in most of our lives. I think the removal of grades would make us feel weird and maybe losing a sense of yardstick, there might be feeling of uncertainty, feeling unsure of how exactly did we do for the assignment we handed in. The grades seem to be a way to give some closure as well sometimes because of the usual routine we are used to. I'm not sure how no grading will affect my performance but i know that the motivation to learn would have to come from somewhere else more internal, like the drive to learn simply because I am interested. I will certainly be less stressful about completing assignments and just work at the assignments as the attitude of a normal homework. |
| I think, at the beginning I would be highly interested in the subject matter. Later on in the semester that intrinsic motivation might be less. That is when the motivation that comes from grading usually kicks in. So, the latter one wouldn't exist. |
| i tend to be lazy when working alone. Lack of grades would take away the goal of scoring higher grades hence i might not perform well. I would work on the studies as much as i like the subject. hence, my performance would depend greatly on my interest on the topics rather the force learning to achieve higher grades. |
| Lack of grading would also result in a decrease in interest and sense of accomplishment. |

Additionally, it is worth mentioning that for some students absence of grading is demotivating and around 50% of the respondents stated that it is important for them to receive the highest possible grade. 71% of respondents stated that grading helps them learn. These answers demonstrate the prioritisation of grades over the learning process because when the course is pass or fail, most students will either pass or fail, there will be no numerical distinction between those who have passed. This data provides an answer to the research question

regarding what students prioritise more: grades or learning and the way they evaluate their own learning and performance. This data supports once again the *Hypothesis 2* regarding students' variations in perception of grading within themselves.

Grades are a source of motivation for students because it gives them a numerical value of how well they have succeeded in their studies. However there were several respondents mention that they would feel less stressed and more willing to try some challenging courses if there were no grading but at the same time they mentioned that they need a certain feedback for their work in order to know which areas could be improved further. The research done by Crooks (1933) demonstrated that the absence of grades would have a positive effect on the learning process as it might help to reveal intrinsic motivations and shift the focus towards the learning itself rather than just receiving the high grade. In the current conditions in education grades serve as a form of feedback for students but they would benefit from having it more elaborate and not just a number or a letter. Similar idea has been mentioned by Zairjanovna et.al. (2022), where the research showed that receiving a feedback-based grading increased the levels of students' engagement, their motivation and academic achievement.

To explore the impact and importance of feedback, the Figure 5 can be used, where the statement regarding the importance of receiving feedback is included. 98% of the respondents agreed with the statement that receiving feedback would help them learn. In the number of the open-ended questions which asked about fair grading, improvements of the grades, necessity of the grades a lot of respondents answered that the grading would be more fair, more informative and more understandable if it would be supported with the verbal or written feedback. Some respondents considered grading essential only because it is the only available tool known to them that can assess their knowledge. This totally supports the hypothesis 3 *“Students tend to perceive feedback together with the grade as an effective method in improving their grading experience and the overall learning outcomes”*. The idea of the beneficial influence of the feedback on student's learning as well as the need for search for the alternatives to traditional grading have been mentioned by a variety of different authors throughout different time periods. Silberstein (1922) and Crooks (1933) shared the idea that the feedback might minimise possible misunderstanding between teachers and students and Carter's study (1952) highlighted that feedback would improve teacher-student interaction, Jansow and Eison (1990) mentioned that meaningful feedback might reduce stress and

improve learning and the study by Brookhart (1999) emphasised the beneficial aspects of giving and receiving feedback even more. Allen (2005) and Feldman (2007) discussed the value of feedback from the point of view of the information that it conveys: it gives more understanding regarding the areas that need to be improved.

From the previous findings, it is possible to conclude that grading, both traditional and alternative, is necessary for the learning process. Another research question aimed to discover whether the students consider grades to be an integral part of their learning. The respondents were asked directly in a form of an open-ended question whether they consider grading essential for their learning. After analysis of the answers, it was discovered that 57% of respondents consider grades as integral part of their learning, supporting their answers with explanations that included motivation, illustration of their knowledge, the only evaluation that exists. 33% of respondents think that grading is not necessary for their learning but they point out the importance of having verbal or written feedback instead of a grade since it provides more insight into how the student can improve further. An important issue that have been mentioned is that it is necessary for the educational system to have numerical grades, as they are taken into consideration when applying for the jobs or applying for a certain programme at university.

So, based on the collected data supported by the researches from the literature overview, it can be concluded that there is a strong interconnection between grades and learning. There are several factors and areas that influence that relation and one of them is the lack of understanding of the purpose of the grades across and within the students, which results in the discrepancies which were presented in this section. The interpretation of the results in this sections allowed to support hypothesis 3, add more supporting arguments to hypothesis 2 and answer the intended research questions.

Psychological impact of grading

Primarily, this topic was not intended to be a part of this thesis, both the questionnaire and the utilised literature was mostly focused on the interconnection between grades and learning and students' perceptions of it. However, during the analysis of the collected data, there has been responses which are worth attention during the research of grading.

Psychological impact of grades has not been a major focus of the researches from the literature overview as well. Undoubtedly, there were mentions of that even in the earlier works, for instance Edgeworth (1888) mentioned that the fact that the emotional state of the students and the elevated stress levels during the exam are not considered during the grading process results in the grade which is not reflecting student's actual level of knowledge. Starch and Elliot (1912) pointed out that grading experience impacts students' overall attitude towards education. Jansow and Edison (1990) in their research imply that grading can cause stress and anxiety which lead to the loss of motivation towards learning. The research by Winter (1993) revealed that the grades create an atmosphere of competition which then shifts the attention from actual learning to the pursuit of the highest possible grade.

However, while some of the researches acknowledged that grades can impact psychological state of the students, they tend to view this aspect more generally rather than personally, meaning that grade on a particular assignment or course can elevate stress and anxiety levels and that is quite situational. But the results of the questionnaire revealed that student can take the grades quite personally.

The previous section confirmed that students consider grading to be an essential part of the learning process and students mentioned that there is a need for the numerical grade in the educational system due to the fact that it is considered when applying for a job or to the university.

This exact aspect of grading leads to the frustration and stress among the students. When asked whether the respondents afraid of receiving the unsatisfactory grade, 60% answered positively. Further, they were also asked to explain the reasons that are behind this fear. Among the answers that explained the worry are the future job and educational perspectives. There has been a noticeable layer of answers that had a deep personal involvement. Table 5, which can be found below, demonstrates some of the answers. Those answers provide an overview of a problem that goes beyond educational setting, that is rooted deeply into the persons' self-image and affects their psychological wellbeing.

The fear of receiving unsatisfactory grade which results in student's loss of self-confidence is the proof that the grading system needs to be improved and adjusted with the consideration of the current realities. Based on the answers, nowadays, the psychological impact is not

Table 5.. The fear of receiving unsatisfactory grade, responses.

| |
|---|
| My self-confidence is fragile and self-doubt strong. I have put my goals very high and got myself in a situation where when I receive an excellent grade I feel relief but anything less feels like a failure. I know its bad, and mostly self-inflected. |
| Actually I'm not sure why, but grades do affect our emotions and communicates in some sense how well we are. Having an unsatisfactory grade is a way of saying we are not good and usually we do not want to face this feeling. Of course, there are also other future considerations. For instance, if I would need grades for future studies opportunities, an unsatisfactory grade could be a stumbling block. |
| My grades have become a reflection of my self-worth. If I get an unsatisfactory score, then its the same as saying that I was accepted into this program as a mistake, that I don't deserve this spot or scholarship and that I stole it from someone better. |
| unsatisfactory grade can have a negative impact on career especially if the student is going for academic career or other career where grades are important. |
| I think about the future and possible employment, for example sometimes during job interviews they want to see grades. So if I have bad grades they'll think I'm lazy. |
| It can never be changed. It stays with you till the rest of your life. It influences your career prospects (top companies hire best students) and your academic prospects (when you apply for new education programs, they check your previous diploma transcripts). Once the damage is done, it stays forever :(It is even bitter if it is unfair. |
| I don't know, maybe because shame is associated with a lower grade |
| It's like failing, not being good enough. Even though grades are not showing my knowledge but rather how I adjust to the system, they still mirror to the society what my worth is |
| Afraid might necessarily a too strong word for it. Scared to be seen as a failure either in my own eyes or someone elses might be a bigger reason for not wanting lower grades. |

addressed properly and thus grading exceeds purely educational setting impacting student's feeling of self-worth.

The complex nature of grading

The previous section highlighted the complexity of the grading and the need for new solution in this area even more, since the impact of the changes in the grading is not purely educational but also personal.

The aim of this subsection is to incorporate the knowledge from the literature overview together with the major topics that were outlined and the results. One of the reasons for doing so is because most of those themes are present in the results of this work which confirms that they remain relevant nowadays. Another reason is the possibility to view those topics multidimensionally and while utilising all the collected data, outline positive and negative aspects that are associated with every topic.

The first theme that has been highlighted a lot in the researches throughout the different time periods and remain relevant nowadays is *subjectivity*, as it was observed in the majority of responses. It has been highlighted when exploring the notion of fairness and how it can be improved. Many respondents mention subjective and unbiased approach is the one that needs to be adopted in order to increase fairness in the evaluation process. Such an approach would

make the grades also more informative in regards with the amount of knowledge received and might increase the *validity and reliability of grades* as well.

It became clear from the responses to the questionnaire that *fairness in grading* remain a crucial issue in the current education as well as it was since the beginning of the 20th century. The connection between *grades and motivation* was the most visible in the questions which revolved around the influence of grades on the learning process and the necessity of grades for the learning, as well as the hypothetical situation of the absence of grades. All of those questions showed that there is a strong correlation between grades and motivation among students and while it serves as an extrinsic motivation, it might be beneficial of thinking regarding other possible ways of motivating students other than grades.

Feedback has been mentioned multiple times by students as a medium for improving learning process and as a tool which might enable better communication and understanding between teachers and students. It is also one of the *alternatives to traditional numerical grading* that students suggest might be beneficial from the perspective of learning. The topic of the *psychological impact of grading* was not the main in this thesis, as has already been mentioned in the section above. Thus the results revealed a deep and profound influence of grades on the students' personalities and even their sense of self-worth. All of the above-mentioned suggests the *complexity of the grading* as it is not only about the evaluation of the student's performance but it entices a much broader meaning and has a strong influence on not only learning but also students as individuals. Judging from the questionnaire, it is possible to assume that improving the grading practices might impact student's wellbeing. Such topics as *standardisation of grading* and *challenges in implementation*, that were formulated in the literature overview section, presented themselves not so significantly as the other topics in this particular study as it was focused on discovering students' perceptions of grading and its impact on learning. Thus these aspects were slightly mentioned by some of the respondents in connection with the topics of necessity of grading and the possible ways for improvement.

Below there will be presented Table 5, that was done on the basis of the theoretical overview as well the analysis of the questionnaire for this study. It will demonstrate positive and negative aspects of the topics that were discussed previously in this work. It is worth highlighting that all those positive and negative aspects have been mentioned by both

researchers of the grading topic and respondents of the questionnaire. The purpose of this table is to link the theoretical knowledge with the practical one acquired by the virtue of the questionnaire and to visualise the complexity of grading and its multi-dimensional nature. Additionally, the summary of positive and negative aspects of the grading based on a particular topic facilitates highlighting the areas that are in need for the improvement.

Table 6. Positive and negative aspects of grading based on the thematic categorisation.

| Theme | Positive aspects of grading | Negative aspects of grading |
|------------------------------------|--|---|
| Subjectivity in grading | <ul style="list-style-type: none"> - provides feedback on student's overall performance; - provides holistic view of student's performance; - teachers tend to assign reliable grades. | <ul style="list-style-type: none"> - leads to inconsistent and biased evaluation; - grades may be influenced by subjective factors such as student's gender; - influenced by personal judgements. |
| Grades and students' motivation | <ul style="list-style-type: none"> - encourages effort; - provides a source of external motivation; - higher expectations lead to increased student's involvement; - encourages goal-setting. | <ul style="list-style-type: none"> - can demotivate if perceived as unfair; - can suppress the internal motivation; - lead to focus solely on grades; - can cause stress, anxiety and reduce overall motivation towards learning. |
| Reliability and Validity of Grades | <ul style="list-style-type: none"> - helps to measure student's progress; - teaches assignment grades based on correlation with academic achievement; - communicates student's achievement. | <ul style="list-style-type: none"> - subject to inconsistency and bias; - emotional state of students is not considered; - inclusion of subjective factors decrease reliability and validity of grades. |
| Complexity of Grading | <ul style="list-style-type: none"> - acknowledges multifaceted nature of assessment; - teachers consider various factors while assessing; - emphasises importance of individuality; - allows comprehensive assessment. | <ul style="list-style-type: none"> - makes grading challenging and variable; - inclusion of multiple factors may complicate grading process; - lack of standardised measurements; - inclusion of multiple factors is challenging in implementation. |
| Standardisation of Grading | <ul style="list-style-type: none"> - promotes fairness and consistency; - standardised tests provide precise measurement of student's achievement. | <ul style="list-style-type: none"> - may oversimplify complex learning; - standardised tests may not capture all aspects of student's learning; - limits individuality and uniqueness of students. |
| Alternatives to Grading | <ul style="list-style-type: none"> - offers potential for more objective assessment; - provides more comprehensive feedback and learning opportunities; - allows for diverse assessment methods. | <ul style="list-style-type: none"> - may lack comparability with traditional grading systems; - may require significant changes in traditional grading practices; - might require teacher training. |

| Theme | Positive aspects of grading | Negative aspects of grading |
|---------------------------------|--|---|
| Fairness in Grading | <ul style="list-style-type: none"> - ensures equal treatment for all students; - ensures equity and transparency in evaluation; - promotes improvement rather than competition; - aims for objective measure of student's performance. | <ul style="list-style-type: none"> - can be still biased and thus lead to unfair grading; - perception of fairness may differ from actual grading standards; - can lead to inequities if not implemented transparently and consistently. |
| Perspectives of Grading | <ul style="list-style-type: none"> - provides diverse viewpoints on assessment; - grades are viewed as a valid measure of student's achievement. | <ul style="list-style-type: none"> - may lead to conflicting approaches resulting in confusion and inconsistency; - subjectivity and biases impact grading outcome. |
| Psychological impact of Grading | <ul style="list-style-type: none"> - provides feedback for improvement; - can motivate to strive for more improvements and success. | <ul style="list-style-type: none"> - can create stress and anxiety for students; - can reduce overall motivation for learning; - can impact student's self-esteem. |
| Challenges in implementation | <ul style="list-style-type: none"> - highlights areas for improvement in grading practices; - emphasises the need for teacher training and professional development. | <ul style="list-style-type: none"> - requires overcoming resistance and changing practices; - can lead to inconsistencies and biased if not implemented correctly. |

Such aspects as subjectivity, impact of grades on the motivation, fairness, reliability of grades and their psychological impact on the students were presented in the most prominent way both in the literature overview and results parts. The impact of grades on the psychological level though has been especially highlighted by respondents of the questionnaire in the open-ended questions but the profound exploration of this begins only in the 21st century. Undoubtedly, some aspects have been mentioned in earlier works but they were focused more on an external factors that increase student's stress levels and anxiety or lead to demotivation (Edgeworth, 1888, Crooks, 1933, Janzow and Eison, 1990, Winter, 1993). But the responses that were presented above in the Table 5 revealed the impact of the grades on a personal level: grades influenced the way students think of themselves and as individuals, as human beings. This aspect has not been researched a lot but based on the results, there is a need for its further research.

The element, the importance of which has been highlighted by both researchers of the different time periods and respondents' opinions, is the feedback that is highly beneficial when provided solely or together with the grade. Feedback is an important element of the grading process when learning is prioritised because it is not enough to just assign a grade to a

student. It just gives a mere image of student's level of knowledge. But when the teacher provides a meaningful feedback, it gives a student a broader image of what could be improved further. However, as feedback is one of the tools in the evaluation process, there is a chance for its misinterpretation and thus it might not be as beneficial for the learning process as mentioned before. The further research on a possible negatives of the feedback should be explored further.

Considering all the above mentioned, the themes together with their benefits and drawbacks, it is possible to conclude that with time as more information and researches are released, the issue of grading is becoming even more complicated as more visions are added and more aspects are researched. But that does not mean that it is senseless to research this topic more. On the contrary, the broader and more complex understanding of the issues is achieved, the more complex methods can be used to deal with these issues.

Even though all the major themes remain relevant nowadays, it cannot be said that throughout the time they remained on the same stage as they were more than 100 years ago. They were developing, alternatives to the traditional grading have been found and the fact that there is a possibility to do the study like this one, overviewing numerous researches, proves that the various aspects of grading have always been in a focus of the researchers in the field of education. The solution or the way for improvement of the evaluation process is still to be found but the foundation for researching that has already been created.

5. Conclusion

Grading is a very complex topic which goes beyond a number or a letter for a particular assignment. It is a combination of numerous factors that affect each other and interconnect with each other. For example, it might seem that the grade reveals actual student's knowledge but when researched more profoundly, as it was observed in the literature overview part and further in the results, it became clear that it is correct only to a certain extent. Nevertheless, grades are an integral part of the educational setting at least nowadays since there has not been any unified solution on what can replace grades.

The literature overview part demonstrated that the topics of, for example, subjectivity of grading or correlation between grades and motivation have its roots in the 19th century in the research of Edgeworth (1888) with the focus on understanding the subjective nature of grades and emphasising the importance of addressing this issue for the purpose of minimising biases during the evaluation process. The researches connected to the subjectivity of grading and its impact on learning continue further by Elliot and Starch (1912), Hulten (1925), Swineford (1947), Collins and Nickel (1974), Brookhart (1993, 1999, 2011), Ismail and Tini (2020) and other researchers that aimed to explore the subjectivity issue deeper. Even though there have been done a lot of research, those issues still remain unresolved, there is no unified solution nowadays regarding how to make grading such that reveals students' performances. Undoubtedly, there are some ways to increase objectivity in the evaluation process and reduce misunderstandings between teachers and students such as inclusion of the feedback together with the grade, or practice combining graded and ungraded assignment, or combining different assessment methods for the purpose of gaining the broadest overview of the student's knowledge. And those aspects have been mentioned not only by the researchers from the literature overview but also by the participants of this study. They emphasised the above mentioned elements as the ones that have a potential to add more fairness to the grading process.

However, there is practically no solution to the subjectivity issue, especially when, for example, grading essays, where it is impossible to avoid opinionated grading because it is a human, a teacher, who grades the piece of work, and consequently some of the teacher's views and opinions naturally influence the perception of the work. However, the researches starting from 19th century, or perhaps even earlier, highlighted the importance of the grading

practices that are chosen since they have a tremendous impact on student's achievement and learning.

Summarising the collected information from the literature overview part and then later analysis of the results, it is yet possible to outline the following suggestions that can be considered in order to improve grading:

1. ***More assessment tools should be used in order to ensure a more objective grading.*** Ebel (1974) and Collins and Nickel (1974) suggested in their researches that it is a system of grading tools that can be improved. It was noticed during the literature overview and then during the analysis of the results that relying only on numerical or letter or pass/fail grades might be rather subjective and might not be beneficial for the learning process. But taking several grading tools, for instance, numerical grade, feedback (written or verbal) and self-evaluation or peer evaluation can bring a broader and less subjective image. The combination of the tools during the evaluation process might be beneficial also based on the participants' responses as they mentioned that the absence of grades can demotivate them. So, in order for such students to be more involved in the process, the combination of various grading techniques might be applied. Additionally, it was pointed out that there is a possibility of decreasing stress levels during the evaluation process if the criteria is clearly communicated and thus this aspects should also be considered as one of the tools for the evaluation.
2. ***Providing the feedback together with the grade.*** Plenty of various researches such as Simon and Bellanca (1976) Abrami et al. (1980), Jansow and Eison (1990), Guskey (2000) Kohn (2011) as well as 98% of respondents of the questionnaire viewed feedback as a solid method for improvement of the grading experience and learning itself. Both in the literature overview and analysis of the results, feedback is mainly portrayed as a constructive tool that enhances learning and motivation. However, it should be noted that feedback is also a form of evaluation so similarly to the traditional grade it might be misinterpreted if it is unstructured or incoherent. Nowadays there is a lack of the research done on the drawbacks or weak sides of the feedback. Consequently, it might be one of the areas for the further exploration and research.

During the analysis of both literature and questionnaire results the above-mentioned points have been highlighted the most by researchers of different time periods and participants of

this study. Based on that it is possible to conclude that there is no unified solution that will fix grading issues but it might be useful to focus on improving different grading tools. For instance, both researches from the literature overview and the analysis of the questionnaire seem to view feedback as a valid method for improving grading experience, increase motivation and encourage learning but as it was stated above, feedback is also a form of evaluation and thus there should be requirements as to how one gives feedback and how ones receives it. It should be highlighted though, that since the hypothesis 2 which is “*There is a variation across and within the students in perception of the grading*” was confirmed, students might not fully acknowledge that feedback is a form of grading and can still misinterpret it in the similar way as the numerical grades. Consequently, more research is needed to be done to explore that further and more profoundly.

Current grading discussions are still revolve around subjectivity of the grades, components that should or should not be considered in the process of grading, reliability of grades. All of those issues have been acknowledged at the end of 19th century. However, the difference is that now there is a broader view, there are more perspectives that has already been explored. The question lies not in the fact of finding the perfect grading scale because most likely such a scale does not and will not exist as long as there are different views and opinions. What is more important in this discussion is to bring the focus to the way teachers convey grades to the students, where emphasis is placed on the learning and improvement and grade is one of the tools to track those and not the ultimate criterion.

It is worth mentioning that subjectivity is mostly viewed as an issue that should be fixed. Students in their responses to the question regarding the ways of increasing fairness in the evaluation process were mentioning the subjectivity as an obstruction to fair grading. Nevertheless, the feedback, that was implied by 98% of the students as a method for improving grading experience, is often subjective and, because it is subjective and conveys teacher’s opinion, it is valuable. So the issue of subjectivity should be viewed as a complex issue which has its benefits as well as drawbacks.

Education is the area where comparison is inevitable: students receive a place at higher educational institutions, they are considered when continuing education further. And it is not a negative aspect. It has its place to be. The negative aspect lies deeper — when students consider grades to be the only valid measure of their knowledge, when students are incapable

of analyse even roughly the extent to which they are knowledgeable in a particular subject. Such an attitude leads to the situation when the curiosity in the subject lasts as long as the course continues. After receiving a passable grade the student considers this subject as a finished one.

So, it is more important to evoke intrinsic motivation in students, make them think of a subject, make them curious about the subject. The grade should not be the motivation or the goal because of its subjective nature. It should be just one of the tools.

6. Discussion

In the conclusion part the results and their connection to researches from the literature overview have been communicated. This section intends to explore the implications of the findings and to give an outline of the areas for further research.

This thesis work revealed a number of important aspects of grading. The fact that even though most of the grading issues have their roots in the 19th century or probably even earlier, they remain relevant and require even more exploration nowadays.

The discussion around the subjective nature of grades, showing how a teacher's personality, impressions, and personal interpretation of terms like "interesting" can lead to variations in grades, has been mentioned by Edgeworth in the research in 1888. Further in 2020 Lipnevich et. al. in their study also explored the subjectivity in grading. Their research showed a great variation in grading criteria between different educational institutions. That means that the same grade in different institutions will mean different things. That brings confusion to numerous areas of education, starting from how students can understand their grade correctly, how parents perceive that grade and what teacher actually meant by that grade. And this discrepancy in understanding can lead to decreased motivation in the subject, that can respectively lead to worsening of the learning outcomes. It is natural that the grading criteria differs depending on the type of the assignment, meaning that it is impossible to grade individual and group work in the same way.

The major problem with that lies in the fact that if the grade varies from teacher to teacher, the validity of this grade is then questionable and in the situation when the teacher changes it, it is then hard to estimate student's actual knowledge and academic performance. The probable solution to this particular subjectivity issue might be in giving clear guidelines into how to grade the specific piece of work so that it is relatively same among different teachers.

Current grading systems are created in a way that they capture only one side of student's performance. For instance, Schneider and Hutt (2014) mentioned that A-F grading scheme has been used for rewarding and punishing. In such a scenario there is very little space for the creativity and novelty and those two components are the core elements for the society to grow and improve the way things work. So, instead of supporting those who are finding new ways of doing something, creating something, current educational systems try to put them back in line. With that in mind, more research should be done on the issue of grading, broader

perspective should be taken into consideration, governmental institutions should be involved in the process because in order to change grading practices, the collaboration between those parties should be present. Most of the current grading practices focus on assigning numerical grades that make it further possible to compare students with the standards and requirements. But as it was mentioned before there are no standardised requirements for the grades and consequently they vary among the educational institutions or even teachers. Standardisation of the grading is an important issue that has been addressed more than a hundred years ago. Hence, the problem is not solved up until nowadays.

There is no arguing with the fact, that there are a lot of different innovative grading practices emerged in the recent years. However, all of them are far from being unified and standardised on the global level. Most of those new practices require professional teacher trainings in order to get acquainted with those practices and become knowledgeable on a way they should be implemented in a classroom. This aspect requires resources and financial support from the governmental institutions or sponsors which takes us back to the point mentioned above that in order to improve grading, there should be a close collaboration between government and educational institutions.

Further Improvements

This thesis work demonstrated the profoundly complex nature of the grading and evaluation. Literature overview part made it possible to notice that the research on how to make grading more beneficial for learning and how to make it more fair and objective has a long and not linear history. The results of the questionnaire revealed students' perspectives towards grading and its impact on learning, motivation and psychological wellbeing. All the above-mentioned proves that the topic still has a lot of open questions that need to be if not answered, then researched deeper in order to uncover more perspectives.

The literature overview section and the results of the questionnaire revealed that feedback is seen as a solution to all of the grading problems. 98% of the participants stated that their learning would benefit from the feedback received from the teacher. Researches by plenty of authors throughout the different time period such as Silberstein (1922), Crooks (1933), Carter (1952), Simon and Bellanca (1976), Abrami et al (1980), Winter (1993), Brookhart (1993, 1999, 2011), Guskey (2000, 2002), Allen (2005), Chan and Luk (2022) and others, suggested

that the feedback is a way to improve the overall grading experience, minimise misunderstandings between teachers and students as well as subjectivity during the evaluation process, increase motivation towards learning etc. However, those researches mention nothing regarding the negative aspects that might be caused by the feedback. And it is possible that those drawbacks exist since feedback, as it was stated in the previous sections, is also a form of evaluation and thus might be perceived differently. Based on that, a broader research on both positive and negative aspects of feedback is needed as well as on the ways it might be perceived by the students and delivered by the teacher.

Moreover, this research was focused on gathering students' perceptions of the grading process and its interconnection with the learning. But further research can explore teachers' perceptions more, as it might reveal the aspects that are unknown to students and the difficulties which teachers are facing during the evaluation process. Later, those two studies can be improved and combined and the broader understanding of the current situation with the grading can be observed.

One more aspect, worth of further investigation is student's subjectivity during the evaluation. For instance, Carter (1952) and Swineford (1947) researched the possible impact of student's gender on the received grade meaning that teacher assigns the grade subjectively considering the student's gender. However, the same subjectivity might be observed between students towards the teacher: student's willingness to learn the subject might be influenced by whether the teacher is female or male. The ways students perceive their grades, the ways they perceive teachers is the area which is understudied and requires more exploration.

The part, which is based on the collected information and in my opinion requires more attention nowadays is the psychological impact of grading on students. This topic has not been researched a lot and thus has a lot of potential because the answers to the open-ended questions of the questionnaire revealed the profound impact of grades on the way students think of themselves as persons. So grades influence not only motivational aspect of the learning but also student's self-worth. Grading goes beyond the educational setting and has a way bigger impact on person's life. Grading affects not only motivation during the studies and creates a habit of a need for validation of person's overall self worth later in a workplace. A person needs to have someone who will give evaluation and without it feels incomplete.

Those aspects highlight the need for the exploration of the grading topic through the prism of its impact on the psychological well-being of the students.

This thesis work revealed a complex and multi-dimensional nature of the grading. This issue has been supported by the studies from various time periods starting from the end of the 19th century and continuing until nowadays. The questionnaire that has been the instrument in this study underpinned the theoretical material that has been found. Regardless of the long history of exploration of the grading issue it still remains essential in the educational setting and requires further research and solutions.

Limitations

This research study, similarly to most of the research studies, inevitably has some limitations even though it was attempted to be designed and executed systematically and scrupulously. Those constraints are crucial to highlight in order to provide more transparency for the reader as well as to identify the areas for further research and improvement.

The first limitation is the sample size, which is relatively small and as a result the generalisation that is made in the results section might not visualise the broad picture. However, for the purpose of being able to make those generalisations trustworthy, the researches on a particular subject will be presented together with the findings.

Another limitation is the participants' demographics. Even though international students have been chosen specifically due to their unique perspectives and experiences from different educational backgrounds, there might be biases in both answering and interpretation that are influenced by cultural differences and socioeconomic status.

The choice of participants is also a limitation not only due to their demographics but also due to the fact that those are students. As it was mentioned in the further improvements the inclusion of the teachers as a participants or researching their perspectives separately will allow to investigate the grading from a wider perspective.

One more limitation is connected to the questionnaire itself because the wording might still be interpreted differently by different people even though the responses seem to capture the essence of the initial question.

In order to mitigate those limitations in the future research, the sample size must be larger and more diverse. Additionally, the use of more than one research method, for instance

questionnaire and interviews, will allow to decrease misinterpretations of the main concepts as well as will provide a better understanding of a cultural background of the participants.

Addressing those limitations is important for a coherent understanding of the grading issue and its need for further exploration to collect and utilise more data for the purpose of making conclusions and generalisations regarding students' experiences more accurate.

References

- Abrami, P. C., Dickens, W. J., Perry, R. P., & Leventhal, L. (1980). Do teacher standards for assigning grades affect student evaluations of instruction?. *Journal of Educational Psychology, 72*(1), 107.
- Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 78*(5), 218-223
- Anderson, L. W. (2018). A Critique of Grading: Policies, Practices, and Technical Matters. *Education policy analysis archives, 26*(49).
- Bowers, A. J. (2009). Reconsidering grades as data for decision making: More than just academic knowledge. *Journal of Educational Administration.*
- Bowers, A. J. (2011). What's in a grade? The multidimensional nature of what teacher-assigned grades assess in high school. *Educational Research and Evaluation, 17*(3), 141-159.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement, 30*(2), 123-142.
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied measurement in Education, 7*(4), 279-301.
- Brookhart, S. M. (1999). Teaching about communicating assessment results and grading. *Educational Measurement: Issues and Practice, 18*(1), 5-13.
- Brookhart, S. M. (2011). Starting the Conversation About Grading. *Educational Leadership*. Retrieved from <https://www.greatschoolspartnership.org/wp-content/uploads/2016/11/Starting-the-Conversation-about-Grading-2.pdf>.
- Burger, R. (2017). Student perceptions of the fairness of grading procedures: a multilevel investigation of the role of the academic environment. *Higher Education, 74*, 301-320.
- Campbell, C. M., & Cabrera, A. F. (2014). Making the mark: Are grades and deep learning related?. *Research in Higher Education, 55*, 494-507.
- Carter, R. S. (1952). How invalid are marks assigned by teachers?. *Journal of Educational Psychology, 43*(4), 218.

- Chan, C. K., & Luk, L. Y. (2022). Going 'grade-free'?—Teachers' and students' perceived value and grading preferences for holistic competency assessment. *Higher Education Research & Development, 41*(3), 647-664.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational assessment, 3*(2), 159-179.
- Collins, J. R., & Nickel, K. N. (1974). A Study of Grading Practices in Institutions of Higher Education.
- Cox, K. B. (2011). Putting classroom grading on the table: A reform in progress. *American Secondary Education, 67*-87.
- Crooks, A. D. (1933). Marks and marking systems: A digest. *The Journal of Educational Research, 27*(4), 259-272.
- Dilman, D., & Smyth, J. (2014). L Melani Christian. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 4th Edition| Wiley.
- Duncan, C. R., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *Alberta Journal of Educational Research, 53*(1).
- Ebel, R. L. (1974). Marks and marking systems. *IEEE transactions on Education, 17*(2), 76-92.
- Edgeworth, F. Y. (1888). The statistics of examinations. *Journal of the Royal Statistical Society, 51*(3), 599-635.
- Eells, W. C. (1930). Reliability of repeated grading of essay type examinations. *Journal of Educational Psychology, 21*(1), 48.
- Farias, G., Farias, C. M., & Fairfield, K. D. (2010). Teacher as judge or partner: The dilemma of grades versus learning. *Journal of Education for Business, 85*(6), 336-342.
- Feldman, K. A. (2007). Identifying exemplary teachers and teaching: Evidence from student ratings. *The scholarship of teaching and learning in higher education: An evidence-based perspective, 93*-143.
- Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: Implications for instruction in measurement. *Educational Measurement: Issues and Practice, 12*(3), 23-30.

- Grimes, T. V. (2010). *Interpreting the meaning of grades: A descriptive analysis of middle school teachers' assessment and grading practices*. Virginia Commonwealth University.
- Guskey, T. R. (2000). Grading policies that work against standards... and how to fix them. *Nassp Bulletin*, 84(620), 20-29.
- Guskey, T. R. (2002). Perspectives on Grading and Reporting: Differences among Teachers, Students, and Parents.
- Hany, K., Proctor, M., Wollenweber, J., & Al-Bataineh, A. (2016). Teacher perception of standards-based grading: Implication and effectiveness. *Journal of Teaching and Education*, 5(1), 749-764.
- Holmes, D. S. (1971). The relationship between expected grades and students' evaluations of their instructors. *Educational and Psychological Measurement*, 31(4), 951-957.
- Hulten, O. E. (1925). The personal element in teachers' marks. *The Journal of Educational Research*, 12(1), 49-55.
- Ismail, I., & Tini, T. (2020). The effect of formative assessment on students' grades in learning English. *MAJESTY JOURNAL*, 2(1), 20-30.
- Janzow, F., & Eison, J. (1990). Grades: Their influence on students and faculty. *New directions for teaching and learning*, 1990(42), 93-102.
- Joan B, G. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of statistics education*, 2(1).
- Knaack, S., Kreuz, A., & Zawlocki, E. (2012). Using Standards-Based Grading to Address Students' Strengths and Weaknesses. *Online Submission*.
- Kohn, A. (1994). Grading: The issue is not how but why. *Educational Leadership*, 52(2), 38-42.
- Kohn, A. (2011). The case against grades. *Educational Leadership*, 69(3), 28-33.
- Kulick, G., & Wright, R. (2008). The Impact of Grading on the Curve: A Simulation Analysis. *International Journal for the Scholarship of Teaching and Learning*, 2(2), n2.
- Lauterbach, C. E. (1928). Some factors affecting teachers' marks. *Journal of Educational Psychology*, 19(4), 266.

- Lee, E., Carberry, A., Diefes-Dux, H., Atwood, S., & Siniawski, M. (2017). Faculty perception before and after implementation of standards-based grading. In *2017 research in engineering education symposium, REES 2017*.
- Lekholm, A. K. (2011). Effects of school characteristics on grades in compulsory school. *Scandinavian journal of educational research, 55*(6), 587-608.
- Lipnevich, A. A., Guskey, T. R., Murano, D. M., & Smith, J. K. (2020). What do grades mean? Variation in grading criteria in American college and university courses. *Assessment in Education: Principles, Policy & Practice, 27*(5), 480-500.
- Marsh, H. W. (1984). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of educational psychology, 76*(5), 707.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International journal of educational research, 11*(3), 253-388.
- Marzano, R. J., & Heflebower, T. (2011). Grades that show what students know. *Educational Leadership, 69*(3), 34-39
- Maurer, T. W. (2006). Cognitive dissonance or revenge? Student grades and course evaluations. *Teaching of Psychology, 33*(3), 176-179.
- Mcbride, A. E., & Kahle, J. (2018). Standards-based grading: An examination of assessment options. *Journal of Education & Social Policy, 5*(4), 140-145.
- McMillan, J. H., & Nash, S. (2000). Teacher Classroom Assessment and Grading Practices Decision Making.
- Meyer, M. (1908). The grading of students. *Science, 28*(712), 243-250.
- Miller, J. J. (2013). A better grading system: Standards-based, student-centered assessment. *English Journal, 111*-118.
- Peters, R., Kruse, J., Buckmiller, T., & Townsley, M. (2017). " It's Just Not Fair!" Making Sense of Secondary Students' Resistance to a Standards-Based Grading. *American Secondary Education, 45*(3), 9-28.
- Randall, J., & Engelhard Jr, G. (2009). Examining teacher grades using Rasch measurement theory. *Journal of Educational Measurement, 46*(1), 1-18.
- Salmons, S. D. (1991). *The relationship between students' grades and their evaluation of instructor performance* (Master's thesis, Radford University).


- Schneider, J., & Hutt, E. (2014). Making the grade: A history of the A–F marking scheme. *Journal of Curriculum Studies*, 46(2), 201-224.
- Scriffiny, P. L. (2008). Standards-based grading. *Educational Leadership*, 66(2), 70-74.
- Silberstein, N. (1922). The variability of teachers' marks. *The English Journal*, 11(7), 414-424.
- Simon, S. B., & Bellanca, J. A. (1976). Degrading the Grading Myths: A Primer of Alternatives to Grades and Marks.
- Smith, J. K. (2003). Reconsidering reliability in classroom assessment and grading. *Educational measurement: Issues and practice*, 22(4), 26-33.
- Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10(2), 176-199.
- Starch, D., & Elliott, E. C. (1912). Reliability of the grading of high-school work in English. *The School Review*, 20(7), 442-457.
- Starch, D. (1913). Reliability and distribution of grades. *Science*, 38(983), 630-636.
- Swineford, F. (1947). Examination of the purported unreliability of teachers' marks. *The Elementary School Journal*, 47(9), 516-521.
- Swinton, O. H. (2010). The effect of effort grading on learning. *Economics of Education Review*, 29(6), 1176-1182.
- Thorsen, C., & Cliffordson, C. (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades. *Educational Research and Evaluation*, 18(2), 153-172. - check out the authors
- Tierney, R. D., Simon, M., & Charland, J. (2011). Being fair: Teachers' interpretations of principles for standards-based grading. *The educational forum* (Vol. 75, No. 3, pp. 210-227). Taylor & Francis Group.
- Voelkel, S., Varga-Atkins, T., & Mello, L. V. (2020). Students tell us what good written feedback looks like. *FEBS Open bio*, 10(5), 692-706.
- Willcoxson, L. E. (2006). "It's not fair!": Assessing the dynamics and resourcing of teamwork. *Journal of Management Education*, 30(6), 798-808.
- Winter, R. (1993). Education or grading? Arguments for a non-subdivided honours degree. *Studies in Higher Education*, 18(3), 363-377.

- Woodruff, D. J., & Ziomek, R. L. (2004). High School Grade Inflation from 1991 to 2003. Research Report Series 2004-04. *ACT Inc.*
- Yesbeck, D. M. (2011). *Grading practices: Teachers' considerations of academic and non-academic factors*. Virginia Commonwealth University.
- Zairjanovna, S. D., Mamadjanovna, K. M., Salaxiddinovna, M. G., & Tulaboevna, T. G. (2022). Grade-Free Learning System as A Means of Creating a Psychologically Comfortable Educational Environment in the Class to Motivate a Child to Successful Learning Activities. *Eurasian Research Bulletin, 14*, 17-20.
- Zoeckler, L. G. (2007). Moral aspects of grading: A study of high school English teachers' perceptions. *American secondary education, 83-102*.

Appendix 1. Questionnaire on students' perceptions of grading



Grading experience

 Mandatory questions are marked with a star (*)

I am Olha Nikolaienko, a Master's degree student in the Education and Learning study program. My study explores students' experiences of grading and the correlation between grades, learning, and motivation.

This is a questionnaire about your grading experience as a student at the university. Some questions might require you to reflect on your previous experience. The main aim of this questionnaire is to understand your perception of the grade during studying and learning.

It will take approximately 20 minutes to answer it.

Your responses are confidential and I will pseudonymize the answers before analyzing the data. This questionnaire does not collect any direct personal information other than the request for an e-mail address (voluntary) in case you are also willing to participate in an interview. This email address will be separated from the answers immediately when the questionnaire is closed. The email addresses are solely used for the purpose of inviting you for the interview and will be deleted after the interview unless agreed otherwise during the interview. Interviews themselves will be pseudonymized during the transcription process.

If you would like to know more about the processing of the data, please see the privacy statement by visiting the [link](#).

Your participation is highly valuable and will help my research a lot.

1. If have read the information about privacy and confidentiality and you agree to participate, mark this below *

I agree

The first set of questions is focused on your **understanding of the grades**. In this case, *understanding* means the knowledge of the content of the grade and the predictability of the grade.

2. Mark to which extent you agree or disagree with each statement (1 - totally agree, 5 - totally disagree) *

| | Totally agree | Agree | Neither agree nor disagree | Disagree | Totally disagree |
|---|-----------------------|-----------------------|----------------------------|-----------------------|-----------------------|
| I know what parameters are included in the grade | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| The grade is an indication of how much I have learned during the course. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Evaluation of my own work usually coincides with the grade that I receive for it. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Grade is a valid measure of my knowledge | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

3. I know why I received a certain grade *

- Always
 Usually
 Rarely
 Never

4. I can ask for clarification from the teacher if I do not understand the reason for receiving a particular grade. *

- Yes
 No

5. What are the reasons that restrain you from asking?

6. Teacher explains the evaluation criteria for the course *

- Always
 Usually
 Rarely
 Never

The second set of questions is focused on your **perception of the grades and evaluation**

7. Have you been in a situation where you were thinking that the grade that you received was lower than you expected? *

- Yes
 No

8. Have you felt that this was unfair? *

- Yes
 No

9. Have you been in a situation where you were thinking that the grade that you received was higher than you expected? *

- Yes
 No

10. Have you felt that this was unfair? *

- Yes
 No

11. You wrote an essay that you really liked. You used good, relevant literature resources and put a lot of effort into it. You are satisfied with the result of your work and expect to receive a good or excellent grade. Finally, the result comes and you receive a satisfactory grade. How would that make you feel? What would be your first thoughts? Would you ask a teacher for a reason for that grade? Why/why not? *

12. You wrote an essay that is OK. You did a decent job but you do not expect to receive a high grade. Finally, the result comes and you receive an excellent grade. How would that make you feel? What would be your first thoughts? Would you ask a teacher for a reason for that grade? Why/why not? *

13. How would you describe fair grading? *

The third group of questions is focusing on the correlation between **grading and learning** and the way how you see this correlation.

14. Do you think that the grading helps you learn? *

- Yes
 No

15. How grades influence your learning process? *

16. Are you afraid of receiving an unsatisfactory grade? *

- Yes
 No

17. In your opinion, why are you afraid of this?

18. Have you been in a situation where you have not taken the course that you were interested to learn only because you were afraid of evaluation? *

- Yes
 No

19. If there were no grades, would you be more willing to take that course? *

- Yes
 No

**20. Mark to which extent you agree or disagree with each statement
(1 - totally agree, 5 - totally disagree) ***

| | Totally agree | Agree | Neither agree nor disagree | Disagree | Totally disagree |
|--|-----------------------|-----------------------|----------------------------|-----------------------|-----------------------|
| When taking a course, I focus on my learning and my improvement rather than on the grades. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I do not put effort during the course which is not graded. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| Receiving the highest grade is the best motivation for me. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| I think, that having feedback together with a grade would be beneficial for my learning | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| It is important for me to receive the highest possible grade. | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

21. Do you think the grading is an essential part of learning? Why/why not? *

22. What is the function of the grades in the learning process, in your opinion? *

23. Imagine, there is no grading throughout the whole semester? How would you feel? How do you think this will influence your performance? *

24. When thinking about how to improve the grading and evaluation process, what suggestions do you have? *