

## Article

# What's in a Grade? Teacher Candidates' Experiences of Grading in Higher Education: A Phenomenographic Study

Cormac McGrath , Ylva Ståhle and Lena Geijer

Department of Education, Stockholm University, 10691 Stockholm, Sweden; Ylva.stahle@edu.su.se (Y.S.); lena.geijer@edu.su.se (L.G.)

\* Correspondence: cormac.mcgrath@edu.su.se

**Abstract:** This study explores teacher candidates' experiences of grading in higher education. A phenomenographic approach was adopted and four qualitatively different categories were identified. Grading was experienced as: self-identification, motivation, personal interpretation and academic enculturation. The results indicate that teacher candidates accept existing grading systems but have difficulty interpreting and explaining them, illustrating areas of importance in teacher education and argues that if teacher candidates do not perceive genuine differences in the performance of assessing by grade descriptors, there is a risk that they may develop an insufficient understanding of grading practices.

**Keywords:** grading; grade descriptors; phenomenography; teacher education



**Citation:** McGrath, C.; Ståhle, Y.; Geijer, L. What's in a Grade? Teacher Candidates' Experiences of Grading in Higher Education: A Phenomenographic Study. *Educ. Sci.* **2021**, *11*, 422. <https://doi.org/10.3390/educsci11080422>

Academic Editor: Eila Jeronen

Received: 28 June 2021

Accepted: 5 August 2021

Published: 11 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

More than 10 years after the introduction of the Bologna agreement in Sweden, the reform has arguably had an impact on teachers' assessment practices as well as students' approaches to learning. Among the main outcomes of the Bologna process (hereafter Bologna) was the implementation of new grading systems. To date, little work has focussed on students' understanding of such grading systems, in particular in the Scandinavian context. This study explores students' understanding of such grading systems, at a university that implemented a seven-tiered grading system as a response to Bologna.

Bologna, when implemented, was intended to be a unifying blueprint for European re-organization of study structures and university degrees. Bologna involves the implementation of a three-cycle higher education system, a mutual recognition of qualifications and learning periods completed at other universities, as well as an implementation of a system of quality assurances in order to strengthen the relevance of learning outcomes in higher education [1]. It continues to be adopted in diverse ways and in adherence to local reform efforts and outcomes. Universities in Sweden, for example, interpret differently how to implement the Bologna agreement, and how to articulate learning outcomes to students. In a similar fashion, universities in Sweden are autonomous in deciding if and how their grading systems are impacted by Bologna. Articulating and communicating about learning outcomes is identified as one of the key elements of the Bologna agreement. The implementation of learning outcomes is seen as a key element from a quality assurance and also a pedagogical perspective [1]. At the same time as Bologna was introduced in Sweden, teacher education in Sweden became increasingly academized [2]. Academisation in this sense involves a move towards academic values and in particular conducting academic research, even for students who will not go on to become researchers, which may lead to a loss of connections to professional practice [2]. One example is the integration of teacher education into the universities, for example the teacher training college (Högskolan för lärutbildning i Stockholm, HLS) in Stockholm, was integrated into Stockholm University in 2008. A consequence of this thrust towards the academization of teacher training is that many educators are academics, without previous teaching experience from schools. Such a

thrust towards the academization may come at the expense of a previous focus on teaching methodology and vocational approaches. Moreover, our own experiences as university teacher educators suggests that few university teachers have formal training in assessing and grading student work, and consequently the knowledge and conceptions of grading practices is often a result of trial-and-error.

Teacher candidates' experiences of Higher Education (HE) studies may have a formative impact on how they engage with their own students in compulsory and upper secondary schools. Not only do teacher candidates' HE studies form the major part of their formal education, but also, the systems of education they encounter are likely to have an impact on how they subsequently act as teachers [3,4]. Research suggests, unsurprisingly, that teacher candidates who have more preparation for teaching are also more confident and successful with students [4]. Preparation in this context, could relate to discussions around their own experiences of assessment and grading. Consequently, this study is concerned with teacher candidates' understanding of grading practices in higher education.

## 2. Learning Outcomes as a Driver of Student Learning

The Bologna agreement has a key pedagogical dimension; that by regulating education with clearly intended learning outcomes, students could best prepare for their studies, understand the expectations involved in studying, and adjust their studies accordingly. One of the major impacts of the implementation of Bologna in the Swedish context was the introduction of such learning outcomes. This view on learning outcomes as one of the most significant factors for successful learning, enabling students to plan and prepare for their studies is reported elsewhere [5].

Establishing learning outcomes and clear assessments may be one way of achieving such clarity, enabling students to gain insights into and influencing their learning processes and trajectories. Other research suggests that, constructively aligned curricula and courses may encourage students to adopt deeper approaches to their studies [6]. At the same time, critical voices suggest that the implementation of learning outcomes in HE offers merely cosmetic value to the discussion of learning in HE [7]. Others still, argue that a top-down implementation of outcome-based education, with clear learning outcomes, may have little effect on teaching practices and may even lead to resistance among the teaching staff (8). This phenomenon of resistance among teaching staffs have been noted particularly as a consequence of the implementation of the Bologna agreement and outcome-based curricula, for example, in Sweden [8,9] where the Bologna agreement, has in some cases only lead to an increased documentarising of education practices.

## 3. Drivers of Learning

Notwithstanding the criticism outlined above, in constructively aligned curricula, assessment and grading are thought to play central roles [10] also for the development of students' study strategies. Moreover, there seems to be broad consensus that assessment practices are fundamental to forming student learning [11–16]. Boud and Molloy [17] argue that assessment and feedback processes in HE are governed by a number of mechanisms: prerequisites, the type of abilities students bring with them to HE settings and their previous experiences of assessment; the curriculum, the kind of assessment practices specified in HE settings and how well the curriculum is designed with assessment in mind; and finally, the lived experience and environment, how well does the HE environment correspond to the curriculum and take advantage of students' previous experiences of assessment [17]. Here, in this study, we are particularly concerned with the students' lived experience of grading practices.

This view of the importance of the learning environment corresponds to other notions on the relationship between teachers' approaches and students' learning strategies whereby teachers, through their practice, have an impact on students' learning and study strategies [3]. The operationalisation of assessment and grading practices requires systematic training of teachers in order to facilitate students' understanding [4–11]. Not only

do students bring different experiences of assessment to HE, they also encounter very different types of assessment and also grading practices in HE [17]. It is likely that teacher candidates have experienced at least one, but very likely several different models and approaches to assessment and grading in HE.

#### 4. Approaches to Assessment Standards

As a way of understanding the different approaches to assessment, Ref. [15] propose a framework to understand practices in HE which consists of four models of assessment: the traditional model, the dominant logic-explicit model, the socio-constructivist model and the community of practice model. Similar to apprenticeship models of learning, the traditional model suggests that assessment practices and criteria are not explicitly articulated to students. Instead, students gradually come to know how standards are set and how quality is assessed. Teachers who adopt this model may take a *laissez-faire* approach to student performance, where judgements are based on tacit standards that are informally communicated in serendipitous ways. The dominant-explicit model is characterised by assessment criteria that articulate standards explicitly, but in a passive way. The dominant-explicit model represents a technocratic view of the learning situation, whereby information is made available to the students, but the information is not processed or made understandable to them [15]. This approach to assessment aligns, to some extent, with a techno-rationalist approach, whereby success of assessment is attributed to clarity of criteria [14]. Such approaches have been criticised for seducing teachers into relying too much on so-called explicit criteria. The dominant-explicit model is based on the premise that clarity is expressed through formal criteria and grade descriptors, implying a shared understanding and consensus among academic staff, as well as students, about how these different assessment artefacts are used and perceived of [13,14].

The socio-constructivist approach acknowledges the importance of joint participation with respect to assessment and grading practices. In line with such an approach, students are actively engaged, becoming familiarised by means of various activities with the assessment criteria, thus creating an understanding of what the criteria mean in practice [15]. It may be reasonable to believe that when the students have access to the criteria, that they are judged by, an opportunity presents itself for them to discuss and argue about their performances. Hence, teachers may need to be more explicit about the motivations underlying their judgements. Ref. [15] argue that it is necessary for university teachers to reflect upon possible ways to interpret criteria, but they also need to look at the range of possible performances related to the same standards and criteria. In this way, clarity gained through assessment criteria must be critiqued by both academic staff and students in a contextual setting, in order to determine if a common understanding exists. Such explicitly shared understanding is unlikely to appear without a systematic approach to assessment and also grading practices [11,18].

O'Donovan et al's fourth model [15] builds on Wenger's theory of community of practice. This approach acknowledges the importance of teachers' and students' mutual engagement in the assessment practices, whereby explicit standards, but also tacit knowledge within the learning community, are discussed and shared to form mutual understandings. Both socio-constructivist and the community of practice approaches align with Orr's argument that assessment is a socio-cultural practice, where understanding of assessment and grading practices is something students and teachers need to reach consensus on [14].

In conclusion to this section, we argue, based on our experiences as teacher educators that the Bologna agreement may have had implications on the Swedish higher education environments. We hold that some of the impact may be predominantly cosmetic, whereby new regulations around learning outcomes and assessment practices may have been reworded but without significant changes in practice. We argue that assessment and grading practices are connected to outcome-based education. For the rest of this paper however, we disconnect assessment and grading practices and focus on the latter. We do this to enable us to focus on students' understanding of grading practices.

We identify that research on students' understanding of assessment and grading is scarce [12,19,20]; research on academics' understanding of grading is also scarce [14,21]; and there is a dearth of empirical research on teacher candidates' experiences of grading. Therefore, this study aims at exploring teacher candidates' experiences of grading in Higher Education.

## 5. Materials and Methods

Phenomenography is a research approach with the aim of describing qualitatively different ways of perceiving, or experiencing different phenomena for example, grading. Phenomenography is based on the basic assumption that people experience phenomena in qualitatively different ways. That is, the same phenomenon can be seen in qualitatively different ways and depending on previous experiences attributed different meanings [22]. Given that people develop a limited number of perceptions of a certain phenomenon, one can therefore find a limited number of qualitatively different categories in a phenomenographic analysis. Phenomenography is explained extensively in the literature [23,24].

In this study, the focus is on teacher candidates' experience of grading practices. To that end, we disconnect assessment and grading here, and argue that they are two separate but interconnected phenomena. By disconnecting we mean that during the interviews we did not probe the students about how assessment was conducted in the different courses or subjects per se.

Data in phenomenographic studies is collected through individual, semi-structured interviews. In this study, data was collected by all three authors. We constructed an interview schedule that could be seen as relatively loosely constructed, but which was firmly structured below the surface. The underlying structure consisted of a number of limited phenomena that we wanted the students to develop their thoughts on during the interview [25,26]. The interview schedule offered a point of departure for the interviews, but more importantly, and in line with the phenomenographic method, the respondents were asked to describe and elaborate on their own experiences of grading. Since the research team included three interviewers, a few open-ended questions were constructed to start with. such as: What is your experience of grading in higher education? How do you process grades and grade descriptors when studying? The role of the interviewer was to ask clarifying and developing questions in relation to the respondents' stories. In total, 13 interviews were conducted, each lasting between 35 and 60 min, all in the interviewers' offices. A test interview was first conducted in order to test how the questions were understood, and subsequently all three authors conducted individual interviews, the recordings were subsequently transcribed verbatim [27].

## 6. Context and Respondents

The context of the study is one of Sweden's largest universities, offering a wide range of programmes at four different faculties, Science, Social Science, Humanities and Law. The university offers four teacher education programmes: Teachers for Upper Secondary School, Vocational Teachers for Upper Secondary School, Teachers for Primary and Secondary School (K-9) and an Early Childhood Programme (ages 1–5). These teacher education programmes are comprised of three parts: subject studies and didactics, educational science studies and practical teaching placements. All four faculties and as many as 30 departments at the university contribute to teacher education programmes. Completed upper secondary school education is the requirement for eligibility to the teacher education programmes. Extracts from criminal records are also required. Competition is low, and almost all applicants are admitted.

The interview respondents were chosen using a purposeful sampling procedure in order to capture a wide variety of experience [28].

Given the focus of describing respondent's variation respondents with different ages and potential experiences were sought after. All 13 participants were attending the final year of a five-year Upper Secondary School Teacher Education Programme (300 ECTS).

They were chosen because they had many years' experience of grading practices in subject studies, didactics and educational science studies and the education included studies in grades and assessment. The respondents studied language or social science-oriented programmes. The respondents had varying experiences regarding previous studies at universities and colleges nationally and internationally, as well as assessment and grading practices and scales at primary and secondary school education as well as tertiary education, see Table 1 for an example of the different subjects and ages represented in the respondents cohort. It was probable that qualitative differences in conceptions could be expressed in this sample and it is common that phenomenographic studies contain, in the region of 10–15 respondents [23]. Other studies will need to be carried out in order to validate or contradict the findings presented here. Our aim was to find as many different conceptions as possible.

**Table 1.** An overview of the respondents.

Subjects	Ages	Gender
Religion		
Swedish		
English		
History	1963	
Spanish	1979	8 women
French	1984, 1985, 1986, 1988, 1992 1993, 1995	5 men
Civics		
Philosophy		
Linguistics		
German		

## 7. Data Analysis

The three researchers individually read the interview transcripts, and then, via the phenomenographic analysis outlined in Table 2 below, constructed the categories presented in the findings section. The researchers repeatedly compared their independently formulated categorizations, where similar ways of understanding the phenomenon were grouped into categories and the most important was articulated, this is a so-called coded reliability control [29]. These categories represent qualitatively different ways of experiencing grading in a distinctive and succinct way. The researchers engaged in a form of negotiated consensus, in which the categorisations were compared, discussed and calibrated before arriving at the final set of categories, so-called dialogical reliability control [29,30]. In the analysis, the transcripts were pooled into one data set. Individual responses were not sought, and the final categories do not represent those of an individual respondent. The results are presented as qualitative and distinctly separate categories, with descriptions and significant quotes that are used to illustrate the meaning of each category. When taken together, the categories constitute the outcome space of the study [22]. Marton and Booth [22] propose the following three criteria for judging the quality of the categories of description developed in a phenomenographic study: the individual categories should each stand in a clear relation to the perception of the phenomenon under investigation, so that each category tells us something distinct about a particular way of experiencing parts of the world; the categories have to stand in a logical relationship with one another, a relationship that can be hierarchical; and the system should be parsimonious, which is to say that as few categories there are should be explicated as feasible and reasonable, for capturing the critical variation in the data.

**Table 2.** Steps in phenomenographic analysis, adapted from Dahlgren and Fallsberg [31].

Steps	Process
Familiarisation	Reading through the interview transcripts to acquire a feel for how the interview proceeded; at this stage all data in the data set is given equal consideration.
Condensation	Identifying meaning units and marking these for the purpose of further scrutiny; the size of the meaning units can vary; different fragments of sentences can be associated with different ways of experiencing the phenomenon.
Comparison	Comparing the units with regard to similarities and differences.
Grouping	Allocating responses that express similar ways of understanding the phenomenon to the same category
Articulating	Capturing the essential meaning of a certain category.
Labelling	Expressing the core meaning of the category; steps 3–6 are repeated in an iterative procedure to make sure that the similarities within and differences between categories are discerned and formulated in a distinct way.
Contrasting	Comparing the categories through a contrastive procedure whereby the categories are described in terms of their individual meanings as well as in terms of what they do not comprise

## 8. Methodological Limitations

This study focuses on one HE institution which is a limitation, but we see this as the first in a number of studies examining teacher candidates' experiences of grading and grade descriptors. By choosing one site we could have a better understanding of the context the respondents shared. Consequently, there may be risks in drawing broader generalisations from the findings. However, half of the Swedish teacher education programmes use the same grading system. We recommend readers to interpret if, and how the findings may be transferable to their own contexts [24]. Phenomenographic studies capture respondents' experiences of phenomena at the time of the interview, and respondents may change their conception of the phenomena, but the categories themselves are thought to be solid over time. Our study was conducted with teacher candidates in social science and the humanities, and this may have had an impact on the outcome.

## 9. Results

In this study, we identify that the students' experiences of grading in higher education are related to two predominant aspects; grades, but also grade descriptors. The following four qualitatively different categories were identified. Grading was experienced as:

- self-identification
- motivation
- personal interpretation
- academic enculturation

Each category is described in more detail below and students' testimonials from the respondents are presented in italics.

### 9.1. Grading as Self-Identification

This category represents how the respondents identify in respect to academic performance and how grading was a part of this process. This category reflects the respondents understanding of how grading, and in particular how grades and grade descriptors reflect their identity as students. The respondents report identifying realistic goals, but also how they experience a discrepancy at times between their own expectations and the outcome of their performances. This meant that the respondents could identify as a C student. Sometimes they might acquire an A, or sometimes a D, but they could not really understand why they got such a grade. The respondents reported that university admission is based on a merit system, and high and low admission grades for education are important instruments for moderating their own self-image, as admission scores have a value in the eyes of society.

One view that was expressed is that the respondents have different personal expectations. In other words, the respondents identify with specific grades, and qualities expressed by the grade descriptors. The respondents stated that the grades categorise different types of students in different ways; they talk about themselves and define themselves on the basis of grade descriptors.

*"I see myself as a C student, I am happy with a C, it shows that I made an effort, but at the same time it doesn't really matter for the kind of work I get later on, and I see no point in putting in the extra effort."*

It also emerged that respondents try to be realistic when assessing the grade, they identify themselves with, so as not to be disappointed with the outcome. Being in a specific grade span was also something that could change from course to course; each position might be a temporary one, depending on the course material, level of engagement.

*"I can see for myself, I can get an E and be satisfied, I can get a D and be satisfied, C is good and B and A are very good. Then you are good."*

*"My grades are generally C and B, and D at some point. I have thought C is good and then you will not be very disappointed or very surprised, no matter what grades you have received."*

*"I have always been realistic with myself, at the same time I have always had high ambitions, but when I got E and D grades repeatedly, I realised it has to do with me".*

Some of the respondents saw themselves as high-performing students; they reported knowing that they could perform better and acquire better grades if they increased their efforts, but they did not acknowledge that high grades had a value in the context of teacher education.

## 9.2. Grading as Motivation

In this category, the respondents reflected on how their own motivation to study was influenced by grading practices. Motivation had both positive and negative dimensions. This category contains different elements of motivation. In one sense, grades are something to strive for, and a multi-tiered grading system increased their motivation and could act as a catalyst for studying. Conversely, grades and grade descriptors could be de-motivating, especially when they were vague, unclear or not properly explained. The respondents reasoned that, in programmes that have tough competition for future employment, students may be more motivated to aim for high grades. However, they reasoned that the future labour market is very good for teachers, so there was no real point in aiming for high grades.

*"I studied law before and then grades were very important because there is such keen competition for the best jobs, but as a teacher you get a job because there are no teachers, so maybe the grades are not a priority for me!"*

*"I don't try to fight for grades but I do a task until I'm satisfied. It doesn't matter that much. I will get a teaching certificate any way."*

The respondents' personal interest in the content of the course influenced their motivation to engage with their studies and motivated them to spend more time on their studies. In this sense, grading was not used by the respondents to inform their studies; instead, they emphasised intrinsic motivation and a desire to know a lot about the subject.

*"I'm trying to get an A. It's pretty unnecessary, but I can't help it. I want to learn everything as well as I can., After all, I will become a teacher."*

It was reported that the seven-tiered scale affects respondents' motivation to perform differently, compared to, for example, a simple Pass-Fail scale. The seven-tiered scale is perceived as putting greater pressure on the students, which stresses some of them, and motivates others. The respondents reported that a Pass-Fail scale was in some sense de-motivating, as they perceived that extra effort did not pay off. Others expressed the opinion that grade descriptors contained cues enabling them to plan their study strategies.

### 9.3. Grading as Personal Interpretation

This category has two dimensions, one related to the respondents' own personal interpretation of grading of a performance and the other related to the experience that their teachers seem to have diverging grading practices and interpret grades and grade descriptors differently concerning a given performance. In other words, there are personal ways to interpret grading. Moreover, the respondents expressed the view that university teachers seem to interpret grades and grade descriptors differently, often describing the same concepts but in different ways. They reported that understanding also means trying to understand a teacher's own, personal interpretation. Interpretation of grades and grade descriptors is not just something particular to the descriptors, per se, but also a feeling one develops.

*"It is always a matter of interpretation. Everyone interprets it in different ways and it depends, well, I guess on the person using it. The teachers seem to have such different ways of interpreting things, one could give a high grade, and the other maybe a fail. That is kind of scary."*

Students reported that they need to be able to de-code teachers' expectations and intentions. They also perceive that what teachers regard as high quality varies, and that teachers attach importance to different things when assessing, and that there was, at times, a misalignment with the grade descriptors. The respondents reported that it is important to attend seminars and lectures to understand what the individual teacher values when grading. Reading the grade descriptors was seldom reported to be enough to grasp an understanding.

*"Usually, what the lecturer or seminar leader says goes beyond the grade descriptors. I read in the beginning, but then I gave up."*

One view expressed was that the grade descriptors are often unclear, e.g., the respondents had difficulty in identifying the qualitative differences between the different grade descriptors, for example between an A and B. The respondents reported that they felt that the teacher is guided by a gut feeling rather than grade descriptors when grading. One view is that teachers set a C grade if they are uncertain about a student's performance. It also emerged that the respondents usually accept a teacher's grading, if it does not appear completely unrealistic and rarely, if ever, lodged formal complaints.

In this way, the respondents reported that it was necessary to have recurring experiences of working with a text or teacher in order to form a sense of what level a grade should correspond to.

*"It can differ between courses what you can judge as an A. You also work up a sense of what corresponds to which grade."*

Respondents reported that when they could not read or acquire a feeling for their teacher's interpretation of the grade descriptors, they were concerned that they could acquire almost any grade on an assignment.

*"It's difficult to know what they expect. For example, this fall I had a course that I read at E level that this was what I would do to pass the course and then it ended up that I got A. How did that happen? I wondered"*

The respondents reported that it is important to understand the relationship between exam questions and grade descriptors, but also expressed uncertainty and reported that they had to develop an interpretation of their own, of what to do when the grades and grade descriptors and exam assignment did not match.

### 9.4. Grading as Academic Enculturation

In this category, grading is understood as a component of academic enculturation, one where students become a close, internal member of the academic community. In order for the respondents to do well in their studies, they reported that they had to understand



what it means to have an academic approach to their studies. Understanding could be made clear through grade descriptors specifically, which served as way of identifying key elements of what a discipline holds important. Having an academic approach was often emphasised during their studies. What is meant by an academic approach is something which is not always explicitly stated by university teachers. Respondents in this category stated that they had "cracked the code" of university studies and reported examples which included being able to write a structured, clear text, using the course literature, where logical argumentation and intellectual autonomy are central requirements. The respondents reported that it is not primarily the content of an essay or exam which determines the grade but often how a text is formulated.

*"We compared an A level text and an E level text, because the information was almost the same. The content may be almost exactly the same but the structure makes one get much better and one much worse grade."*

Academic writing in particular, was described as something that takes time to learn and is contextual in nature, so the respondents perceived there were differences at the different departments they attended. It was reported that knowing the environment and discipline involves a gradual process of understanding how academics speak about expectations, and that it takes time to learn to identify what may be concealed behind the grade descriptors.

*"Things were most confusing at the Department of Education, it was like always, it depends on this or that. My experience from language studies is that things are much clearer there."*

*"I had studied at university before, so maybe it was faster for me. I already had that academic writing experience, so I had an advantage over many other very young fellow students."*

*"In the beginning, it was difficult to know what is a C answer, what is an A answer, what is well developed and nuanced and so on . . . Now that I am in the last term I have developed a sense of what it means."*

Understanding what an academic approach is, is achieved in many ways, among other things, through comments from teachers, analysis of these comments and applying the insights gained when writing new texts. Respondents stated that language skills are important. In order to do well, you must be able to write rich, nuanced texts based on academic literature, but also decipher the meaning of the grade descriptors. The respondents reported that grades and grade descriptors are presented in a language that is not readily available to the students when they begin their academic careers, for example the concept of problematisation is a concept that students find difficult. Experiences of being assessed helped the respondents create an understanding of what the concepts mean. A recurring view was that grade descriptors are formulated in approximately the same way for all courses and are not specifically indicative of individual courses. Some of the respondents reported that this misalignment leads them to stop reading the grade descriptors towards the end of their education, but this seems to matter less as they had begun to understand what is expected of them in academia.

## 10. Discussion

This study aims at exploring teacher candidates' experiences of grading in higher education. A phenomenographic approach is used. Four categories are identified: grading, as self-identification, as motivation, as personal interpretation and as academic enculturation. We identify two dominant aspects, grades and grade descriptors.

The categories of self-identification and motivation represent somewhat more inclusive ways of understanding grading practices. Respondents reported that they self-identified with a specific grade, that they had always been a C or an A student. At the same time, it was reported that the grades themselves were not a key element of motivation, that they

wanted to feel satisfied with themselves and their own performances. It is not possible to ascertain if this is a specific outcome of a study conducted on teacher candidates, for example, research elsewhere has shown that students in other subjects are much more grade sensitive, and in such disciplines grade descriptors may also play a more important role [32]. More work is needed to capture and understand the value grades play in teacher education better. Previous studies also suggest that, in particular, grades are understood as a measure of self-esteem which aligns with the category motivation [33].

At the same time the respondents did not seem to be entirely motivated by the grades themselves, nor did they use the grade descriptors to guide their studies. One possible interpretation is that the students are driven by intrinsic motivation coupled with an awareness that high grades had little value when students applied for work as teachers.

In the category grade descriptors as, personal interpretation we could see that many students had difficulty in differentiating between the different tiers, so that the difference between an A grade and a B grade when articulated in the grade descriptors was seldom clear and was often understood as a slight semantic difference. Similar findings, albeit in other contexts, have been suggested elsewhere, illustrating how students understand assessment practices [16,19,33–35]. We argue that if students do not perceive genuine differences in the grading of performances, there is a risk that they may develop insufficient conceptualisations of grading practices, where the idea of fairness and student understanding is subservient to the idea of public authority regulation. We postulate that this may be particularly troubling when it comes to teacher candidates who will spend a considerable amount of their professional practice grading, and who also act as gatekeepers for students in terms of professional advancement. While there seems to be a consensus in the literature of the value of using grade descriptors [18,19,36], there are also risks involved in viewing explicit descriptors without facilitating the students' understanding of what they represent and how they might be achieved [13,15].

In this study, we could detect elements of both positive and negative motivational aspects in regards to grades and grade descriptors; this could be seen in the categories self-identification, and motivation. If grades and grade descriptors are unable to express distinct differences, we fear that the value of the grading may be diminished in the eyes of students in general and teacher candidates in particular.

Both categories personal interpretation and academic enculturation also represent more inclusive ways of understanding grading. In personal interpretation there are multiple ways of interpreting grade descriptors. A personal understanding is something the respondents said may develop over time, but they also reported trying to understand how teachers may have interpreted the same phenomena. This shift in thinking from a less to more inclusive understanding of grading is, we argue, an important step for teacher candidates to make. This finding also aligns with previous studies in relation to understanding assessment practices [15], but also previous work conducted on grade descriptors [4,16,20]. In this study, the category academic enculturation is the most inclusive way of understanding grading as it places the respondent in a socio-cultural context of academia, where the value of a performance that corresponds to a certain grade is something that is decided upon within a specific community, department or discipline, and where the grade descriptors have contextual meaning.

The background of this paper is two-pronged, in part, we lean on the ongoing academization of vocational and, in this case, specifically teacher training. Moreover, the paper also frames the study in a context still coming to terms with a move towards Bologna harmonisation [1]. Consequently, this study aims at exploring teacher candidates' experiences of grading in higher education in an environment that is adjusting to Bologna, where teacher trainers are to a high degree university academics who lack formal teacher training. In light of our findings, one concern is that the university in question has adopted a cosmetic form of outcome-based education, as a way to comply with the Bologna agreement. However, our concern is that the same university has struggled to implement a criterion referenced grading, where a common grading system is adopted and where grade descriptors are

used to articulate differences in performance. In light of this, it may be more prudent that grading systems are decided upon at faculty or departmental level and not, in themselves be seen as a way to align to Bologna. That would mean that grades might better represent the disciplines, or the department understanding of how knowledge is best understood, articulated and graded.

We postulate that this would also make it more intuitive for teachers to share exemplars. In this study all the elements of an outcome based and criterion referenced grading are in place, but the findings suggest that grading practices do not act as a way of facilitating student learning. In other words, the environment may not live up to the intentions stated in the curriculum [17,18]. Another possible interpretation is that students, but also teacher educators need to be aware that assessment and grading practices are always contextual, always a socio-cultural phenomenon, where students and academic staff need to arrive at a common understanding [19]. In the literature, some scholars have identified using exemplars as a way of facilitating student understanding [16,35]. In a similar way, descriptors could be used to discuss ways to shape conversations around student learning where the focus of dialogue could be on internalising, but also problematising what excellence is and how it develops in the respective disciplines [13]. We concur with those findings and argue that exemplars would not only be a way of acknowledging the personal interpretations of the university teachers and the students but would also go some way to demonstrating how disciplinarians conceptualize knowledge and, consequently, may act as part of the process of academic enculturation in an increasingly academized teacher education context. Moreover, we argue that this process of developing and understanding of grading practices is a key part of teacher education.

## 11. Conclusions

This study demonstrates teacher candidates' understanding of grading practices in higher education, suggesting that while teacher candidates accept the existing grading system, they have difficulty in explaining its finer points. As we point out in the limitations section, this paper has a small sample and caution should be applied when considering the results. However, as teacher educators ourselves, we are concerned that student candidates are, at times, unable to discern the difference between a different grade descriptors and grades, and fear that might have an impact on their own future practice as teachers. We identify a number of implications; in part, teacher education should be aware of such tendencies among students and also form conversations around grading and grade descriptors. We conclude that teacher educators in higher education settings, who lack formal teacher training, but who are charged with the task of teaching future teachers might benefit from discussing and framing assessment practices in Higher Education. This may enhance teacher candidates understanding of the role assessment plays in learning, but moreover also what facilitates their transition into the teacher occupation.

**Author Contributions:** Conceptualization, C.M.; writing—original draft, C.M., Y.S. and L.G. data collection, C.M., Y.S. and L.G., data Analysis, C.M., Y.S. and L.G., writing—review and editing, C.M., Y.S. and L.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** Not applicable.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Respondents were not asked to consent to their data being publicly shared and therefore the data is not available in its current state.

**Acknowledgments:** Our thanks to the anonymous reviewers for their insightful comments.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Keeling, R. The Bologna Process and the Lisbon Research Agenda: The European Commission's expanding role in higher education discourse. *Eur. J. Educ.* **2006**, *41*, 203–223. [CrossRef]
2. Magnell, M.; Geschwind, L. A seamless blend of research and professional practice: Dual coupling in engineering education. *High. Educ. Res. Dev.* **2019**, *38*, 807–818. [CrossRef]
3. Trigwell, K.; Prosser, M.; Waterhouse, F. Relations between teachers' approaches to teaching and students' approaches to learning. *High. Educ.* **1999**, *37*, 57–70. [CrossRef]
4. Darling-Hammond, L. How teacher education matters. *J. Teach. Educ.* **2000**, *51*, 166–173. [CrossRef]
5. Hattie, J. *Visible Learning for Teachers: Maximizing Impact on Learning*; Routledge: London, UK, 2012.
6. Wang, X.; Su, Y.; Cheung, S.; Wong, E.; Kwong, T. An exploration of Biggs' constructive alignment in course design and its impact on students' learning approaches. *Assess. Eval. High. Educ.* **2013**, *38*, 477–491. [CrossRef]
7. Brancaleone, D.; O'Brien, S. Educational commodification and the (economic) sign value of learning outcomes. *Br. J. Sociol. Educ.* **2011**, *32*, 501–519. [CrossRef]
8. Baldwin, R. *Changing Practice by Reform. The Recontextualisation of the Bologna Process in Teacher Education*; Acta Universitatis Gothoburgensis: Gothenburg, Sweden, 2013.
9. Barman, L.; Bolander-Laksov, K.; Silén, C. Policy enacted—teachers' approaches to an outcome-based framework for course design. *Teach. High. Educ.* **2014**, *19*, 735–746. [CrossRef]
10. Biggs, J.B. *Teaching for Quality Learning at University: What the Student Does*, 2nd ed.; Society for Research into Higher Education; Open University Press: Philadelphia, PA, USA, 2003; p. 309.
11. Sadler, D.R. Specifying and promulgating achievement standards. *Oxford Rev. Educ.* **1987**, *13*, 191–209. [CrossRef]
12. Sambell, K.; McDowell, L. The construction of the hidden curriculum: Messages and meanings in the assessment of student learning. *Assess. Eval. High. Educ.* **1998**, *23*, 391–402. [CrossRef]
13. Norton, L. Using assessment criteria as learning criteria: A case study in psychology. *Assess. Eval. High. Educ.* **2004**, *29*, 687–702. [CrossRef]
14. Orr, S. Assessment moderation: Constructing the marks and constructing the students. *Assess. Eval. High. Educ.* **2007**, *32*, 645–656. [CrossRef]
15. O'Donovan, B.; Price, M.; Rust, C. Developing student understanding of assessment standards: A nested hierarchy of approaches. *Teach. High. Educ.* **2008**, *13*, 205–217. [CrossRef]
16. Hendry, G.D.; Armstrong, S.; Bromberger, N. Implementing standards-based assessment effectively: Incorporating discussion of exemplars into classroom teaching. *Assess. Eval. High. Educ.* **2012**, *37*, 149–161. [CrossRef]
17. Boud, D.; Molloy, E. Rethinking models of feedback for learning: The challenge of design. *Assess. Eval. High. Educ.* **2013**, *38*, 698–712. [CrossRef]
18. Sadler, D.R. Interpretations of criteria-based assessment and grading in higher education. *Assess. Eval. High. Educ.* **2005**, *30*, 175–194. [CrossRef]
19. O'Donovan, B.; Price, M.; Rust, C. The student experience of criterion-referenced assessment (through the introduction of a common criteria assessment grid). *Innov. Educ. Teach. Int.* **2001**, *38*, 74–85. [CrossRef]
20. Bell, A.; Mladenovic, R.; Price, M. Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars. *Assess. Eval. High. Educ.* **2013**, *38*, 769–788. [CrossRef]
21. Tan, K.H.K.; Prosser, M. Qualitatively different ways of differentiating student achievement: A phenomenographic study of academics' conceptions of grade descriptors. *Assess. Eval. High. Educ.* **2004**, *29*, 267–282. [CrossRef]
22. Booth, S.; Marton, F. *Learning and Awareness*; Lawrence Erlbaum: New York, NY, USA; Mahwah, NJ, USA, 1997.
23. Marton, F. Phenomenography—describing conceptions of the world around us. *Instr. Sci.* **1981**, *10*, 177–200. [CrossRef]
24. Bowden, J.A.; Walsh, E. *Phenomenography*; RMIT University Press: Melbourne, Australia, 2000; Available online: <https://search.informit.org/doi/book/10.3316/informit.0864590199> (accessed on 28 May 2021).
25. Prosser, M. Using Phenomenographic Research Methodology in the Context of Research in Teaching and Learning. In *Phenomenography*; RMIT University Press: Melbourne, Australia, 2000; pp. 34–46. Available online: [https://scholar.google.se/citations?view\\_op=view\\_citation&hl=th&user=RbWEe70AAAAJ&start=20&pagesize=80&sortby=pubdate&citation\\_for\\_view=RbWEe70AAAAJ:HDshCWvjkbEC](https://scholar.google.se/citations?view_op=view_citation&hl=th&user=RbWEe70AAAAJ&start=20&pagesize=80&sortby=pubdate&citation_for_view=RbWEe70AAAAJ:HDshCWvjkbEC) (accessed on 28 May 2021).
26. Larsson, S. *Kvalitativ Analys-Exemplet Fenomenografi*; Studentlitteratur: Lund, Sweden, 1986.
27. McGrath, C.; Palmgren, P.J.; Liljedahl, M. Twelve tips for conducting qualitative research interviews. *Med. Teach.* **2019**, *41*, 1002–1006. [CrossRef]
28. Patton, M.Q. *Qualitative Evaluation and Research Methods*; SAGE Publications, Inc.: Thousand Oaks, CA, USA, 1990.
29. Åkerlind, G.S. Variation and commonality in phenomenographic research methods. *High. Educ. Res. Dev.* **2005**, *24*, 321–334. [CrossRef]
30. McGrath, C.; Stenfors-Hayes, T.; Roxå, T.; Laksov, K.B. Exploring dimensions of change: The case of MOOC conceptions. *Int. J. Acad. Dev.* **2017**, *22*, 257–269. [CrossRef]
31. Dahlgren, L.O.; Fallsberg, M. Phenomenography as a qualitative approach in social pharmacy research. *J. Soc. Adm. Pharm.* **1991**, *8*, 150–156.

32. Öhrstedt, M.; Scheja, M. Targeting efficient studying—first-semester psychology students' experiences. *Educ Res.* **2018**, *60*, 80–96. [[CrossRef](#)]
33. Hawe, E. "It's Pretty Difficult to Fail": The reluctance of lecturers to award a failing grade. *Assess. Eval. High. Educ.* **2003**, *28*, 371–382. [[CrossRef](#)]
34. Carless, D. Differing perceptions in the feedback process. *Stud. High. Educ.* **2006**, *31*, 219–233. [[CrossRef](#)]
35. Handley, K.; Williams, L. From copying to learning: Using exemplars to engage students with assessment criteria and feedback. *Assess. Eval. High. Educ.* **2011**, *36*, 95–108. [[CrossRef](#)]
36. Rust, C.; Price, M.; O'Donovan, B. Improving students' learning by developing their understanding of assessment criteria and processes. *Assess. Eval. High. Educ.* **2003**, *28*, 147–164. [[CrossRef](#)]