

# Qualitative Assessment Guidelines for MSc Graduation Projects in Electrical Engineering

Georgi Radulov

Electrical Engineering (EE) faculty,  
Eindhoven University of Technology  
Eindhoven, the Netherlands  
[G.Radulov@tue.nl](mailto:G.Radulov@tue.nl)

Ramiro Serra

Electrical Engineering (EE) faculty,  
Eindhoven University of Technology  
Eindhoven, the Netherlands  
[R.Serra@tue.nl](mailto:R.Serra@tue.nl)

Twan Basten

Electrical Engineering (EE) faculty,  
Eindhoven University of Technology  
Eindhoven, the Netherlands  
[A.A.Basten@tue.nl](mailto:A.A.Basten@tue.nl)

Cecilia Martinez

Universidad Nacional de Córdoba-  
CONICET  
Córdoba, Argentina  
[cecimart@gmail.com](mailto:cecimart@gmail.com)

Martine Greijmans

Electrical Engineering (EE) faculty,  
Eindhoven University of Technology  
Eindhoven, the Netherlands  
[M.H.M.Greijmans@tue.nl](mailto:M.H.M.Greijmans@tue.nl)

**Abstract**—This paper presents novel qualitative assessment guidelines for grading MSc graduation projects in the Electrical Engineering (EE) Faculty of the Eindhoven University of Technology (TU/e). The goals of the guidelines are to provide a grading system that is impartial, time-efficient, consistent, reproducible, transparent, and conducive to enhancing the learning experience. Furthermore, the guidelines establish a taxonomy for evaluating final projects, creating a shared language and a set of criteria across the extensive and varied faculty of EE. The grading is based on five categories: student-centric, work-centric, process-centric, documentation-centric, and communication-centric aspects, where each category features four aspects. For each category and every aspect of it, this paper provides a qualitative scale, totaling in 24 different qualitative gradients.

**Keywords**—assessment, final projects, taxonomy, grading.

## I. INTRODUCTION

Quality of assessment practices within higher education institutions have been a focus of concern in the last decade [1]. The need to address external accreditation agencies' standards [2], to improve the quality of assessment practices that benefit student learning, and to generate greater consistency of grading [3] have required institutions to revise their assessment practices [4]. In particular, many higher educational institutions face the challenge of developing shared assessment criteria to evaluate final projects such as master-level theses. Some studies have found a wide variation in school assessors' dissertation assessment [4] and lack of professors' understanding of evaluation assessment criteria [5]. Grading tools seem to be an important element of the assessment process since they provide definitions that contribute to accuracy, validity and consistency [6]. Poorly designed grading tools present major challenges impacting grade integrity [7]. In spite of the recognized value of grading tools, previous studies recognize that analyzing the process of creation of grading tools is needed to understand their capacity to improve student learning. Reviewing twenty studies on grading tools, Reddy and Andrade [8] found that in general, professors did not describe the criteria and the assumptions behind the process development.

Grades mean different things to different people, at different moments. This may form a problem with respect to the quality assurance of education. Therefore, it is necessary to explicitly define the assessment system and the meaning of the terms used in such a system. Recognizing as well that every grading system presents tensions such as: a) the level

of detail in the rubrics vs the applicability of such rubrics; b) the thorough and in-depth assessment of every aspect vs time efficiency; c) the level of independence of assessors and non-experts opinions vs the supervisor(s) of the project being assessed. All in all, seeking consensus requires the involvement of a large number of faculty members to define common and shared criteria. This trades off with time-efficiency and entails the risk of not finding any consensus at all, and hence creates a need for a common, transparent, reference for the grading scale.

Regarding these concerns, the Examination Committee of the Faculty of Electrical Engineering at Eindhoven University undertook the task of developing guidelines for improving quality and consistency of master thesis assessment. In line with Grainger proposing to engage a group of faculty members to enhance the assessment process [6, 9], a special commission at TU/e analyzed, defined and described assessment criteria and achievement standards. This paper contributes to the line of research that examines the process of grading tool design in higher education [6, 8, 10].

## II. THE CHALLENGE

### A. Aims of the assessment guidelines

Since the Bloom taxonomy in 1956 [11], there has been a strong consensus that achievement standards should focus on the intellectual quality of students' academic work that recalls and recognize “knowledge and the development of intellectual abilities and skills” [11]. Developing both criteria and achievement standards that describe in detail the intellectual quality of academic work requires developing a common understanding of educational goals and student levels of achievement [12]. This section describes the larger aims guiding the assessment guidelines development for a master thesis. The committee agreed that an efficient grading tool [6] offers criteria and standards that allow grade fairness, time-efficiency and grade uniformity: **Grade fairness** is based on two principles [13]. The first principle is impartiality and consistency. Consistency is needed particularly with respect to student insertion in society and industry or academia. The second aspect is that a fair grade should be based on student competence. Fair grades will not provide the students a feeling of expertise, proficiency or mastery where none of these have been actually achieved. A fair grade is defined, as much as possible, with respect to absolute criteria and avoids relative definitions such as the ones brought about by “curve-fitting” grading. **Time-efficiency** is what helps to make a solid and sound grading feasible in a scalable, growing educational

institution. The use of time is relevant in a context where professors and evaluators have too many academic demands [14]. A thorough grading process might be associated with an increased time for grading.

**Grade uniformity** in higher education has been a concern in the past decade [15, 16]. Uniformity assures that throughout the whole department, the interpretation given to grades does not vary from examiner to examiner, from department to department, or from year to year. A uniform grading system is one which is independent of the specific topic, discipline, supervisor, or panel, and it solely depends, in essence, on the level of mastery reached by a student under assessment.

**Reproducibility and repeatability** are linked to grade uniformity. In essence, a reproducible grading process would assign the same grade to the same student (the same final results) when assessed by different evaluation committees. A repeatable grading process assures that an evaluation committee would assign the same grade to different students who achieved the same level.

Grading should not only be a classification process where students are rewarded or penalized. It should be an integral part of the learning process. Grading that is **conducive to learning** is present throughout the whole project development and helps students and teachers reach the desired objectives [17, 18]. Grading thus becomes a distributed process, and not something that comes at the end of the graduation project, where little or no reflection is possible. A grading system that is conducive to learning helps teachers to be able to communicate more effectively (beforehand, with increased clarity, etc.) to students about the criteria and culture behind the grades.

### B. Motivation and Objectives

Defining a grading system like the one described in Sec. II.A is also related to a university's governance and leadership in higher education. Universities should self-regulate their academic standards and, in this sense, define their believed optimal grading system. In this paper, we offer guidelines and qualitative grading scales that can be the basis for tailored modifications with respect to the particular needs of the educational programs. Instead of using the popular quantitative scale: *Sufficient-Average-Good-Very Good-Excellent*, which requires an external reference, we propose to assign qualities to the different grades [4].

We also feel motivated by a need to support and help evaluation committees. In this sense, it is crucial to clarify the role of the supervisor(s) and the assessors. The process of assessment is more productive when the group forms a team, a committee, working towards a common goal, avoiding antagonizing or polarizing the discussion. We acknowledge that grading is a cultural construction. Highly heterogeneous evaluation committees usually encounter difficulties in finding consensus for grades [7]. Our proposal delivers guidelines to support evaluation committees. Thus, our guidelines provide a common language, a taxonomy, for the grading process.

Explicitly written guidelines can assist our young and still inexperienced colleagues. In our evaluation committees, there are often people with limited experience in examination, e.g., people from the industry or junior staff members, who need some definitions and guidelines on how to translate their own

perception on the results of the projects and the knowledge of the student into the grading scale from 1 to 10.

Common guidelines set a common reference that can prevent that different supervisors, departments, etc., deviate in their project evaluations significantly, while the faculty still issues a common diploma [3, 4, 6, 9]. The guidelines proposed in this paper are abstract, flexible and require personal interpretation with respect to the specific projects in order to account for why a specific grade is given: 10, 9, 8, etc. The guidelines should assist and structure the discussions in our evaluation committees. The committees are usually composed of people with different expertise who should align on a common grade. These guidelines provide a common reference for discussions and interpretations.

The shared grading guidelines should reduce bureaucracy and limit workload. They may allow for reducing the number of evaluators in committees. Supervision may in part be delegated to junior academics, including PhD students, while maintaining the quality of supervision. The provided guidelines should not require additional paperwork. They should help examiners formulate motivation for their grades.

Last but not least, the guidelines are meant to help students and assist their process of learning during projects. The guidelines motivate students by providing grade definitions that will clarify the goals and learning objectives for the student in the project. The student will have an a-priori estimation about her/his own levels, which will avoid big misalignments between the expectations and actual results. Moreover, the student is expected to improve the quality of her/his work in the process of meeting the guidelines. Furthermore, the guidelines make the student learn and strive for a higher performance. In the process of preparation for intermediate and final assessments, the student will consult the guidelines, think about these, and learn from these.

Considering the aims and motivation of our work, our main objectives were to define, implement and assure a grading system that: a) is fair, time-efficient, uniform, reproducible and repeatable, and conducive to learning; b) helps evaluation committees assess student learning and performance through clear and meaningful criteria; c) fosters the dialogue that surrounds grades and defines their meaning to various audiences: faculty, university, society, industry, government, press, accreditation committees, etc.

## III. MAIN QUALITATIVE SCALE

The Dutch grading scale is from 1 to 10, with 6 being the passing level. Based on this premise, our first task was assigning specific qualities to the passing grades, so that assessors can test grading ideas against different qualities of a category being assessed. TABLE I summarizes the main qualitative scale. The base level is 8, called "*Proficient*" and the assessment process starts there. If there are shortfalls, then the lower qualities "*Explicit*" (7) and "*Implicit*" (6) should be considered. If the project makes clear contributions to the field, then the higher qualities "*Innovative*" (9) and "*Visionary*" (10) should be considered. The qualitative scale is abstract, and consequently needs to adapt to the particular category being discussed. However, even in such an abstract form, we can define qualities associated with the treads. Level 6, "*Implicit*", expects results, which the student may have not interpreted adequately. Level 7, "*Explicit*", expects that the assessed results are demonstrated, and their limits are known to the student. Level 8, "*Proficient*", expects that the assessed

results and performance are according to the expectations by the time of project definition. Level 9, “*Innovative*”, expects a level of novelty and results that exceed expectations. The highest level 10, “*Visionary*”, expects clear utilization directions or scientific impact of the presented novelties.

TABLE I. MAIN QUALITATIVE SCALE

| Grade | Qualitative scale |                                   |                           |
|-------|-------------------|-----------------------------------|---------------------------|
|       | Tread             | Performance                       | Notes                     |
| 10    | <i>Visionary</i>  | Exceeds expectations              | Breakthroughs are present |
| 9     | <i>Innovative</i> |                                   | Novelties are present     |
| 8     | <i>Proficient</i> | According to all expectations     |                           |
| 7     | <i>Explicit</i>   | Falls shorts in some expectations | Results are demonstrated  |
| 6     | <i>Implicit</i>   |                                   | Results are present       |

#### IV. ASSESSMENT CATEGORIES AND THEIR ASPECTS

The committee established that graduation projects at TU/e were assessed from five different points of view, which we called “categories of evaluation”. Each of these five categories specifies a further set of four aspects, in the form of rubrics, which relate to important characteristics and substantiate the assessment process. TABLE II summarizes the five different categories for evaluation together with their aspects.

TABLE II. MAIN QUALITATIVE SCALE

| Category                 | Aspects                                                                       |                                                          |
|--------------------------|-------------------------------------------------------------------------------|----------------------------------------------------------|
|                          | Rubrics                                                                       | Notes                                                    |
| Specialization           | Quality of literature review                                                  | Student-centric<br>It is about the student               |
|                          | Level of specialized knowledge                                                |                                                          |
|                          | Disciplinary knowledge                                                        |                                                          |
| Research and Design      | Ability to connect problem definition to research field or sub-questions      | Work-centric<br>It is about the deliverables and results |
|                          | Formulation of research questions                                             |                                                          |
|                          | Quality and quantity of established results                                   |                                                          |
|                          | Creativity, originality, innovative value                                     |                                                          |
| Execution                | Critical attitude towards results, methods, scope and perspective of research | Process-centric                                          |
|                          | Level of independence                                                         |                                                          |
|                          | Commitment and dedication                                                     |                                                          |
|                          | Time planning                                                                 |                                                          |
| Report                   | Effectiveness                                                                 | Documentation-centric                                    |
|                          | Readability of report                                                         |                                                          |
|                          | Problem formulation                                                           |                                                          |
|                          | Quality of content                                                            |                                                          |
| Presentation and Defense | Structure and organization of report                                          | Communication-centric                                    |
|                          | Coverage of research outcomes                                                 |                                                          |
|                          | Presentation skills                                                           |                                                          |
|                          | Quality of supporting material                                                |                                                          |
|                          | Discussion skills                                                             |                                                          |

The five categories capture five relevant points of view on a given project:

- Specialization* is a category which concentrates on the student, i.e., student-centric;
- Research and Design* concentrates on the results, i.e., work-centric;
- Execution* concentrates on how the project is carried out, i.e., process-centric;
- Report* concentrates on the written communication in the project, i.e., documentation-centric;

e) *Presentation and Defense* concentrates on the interactive dissemination of knowledge and results, i.e., communication-centric.

#### A. Categories

The main qualitative scale applies to the five categories and to their aspects. The gradient adapts according to the needs of the assessment. Figure 1 shows an infographic of the high-level gradients for the five main categories, explained and presented in a textual form below.

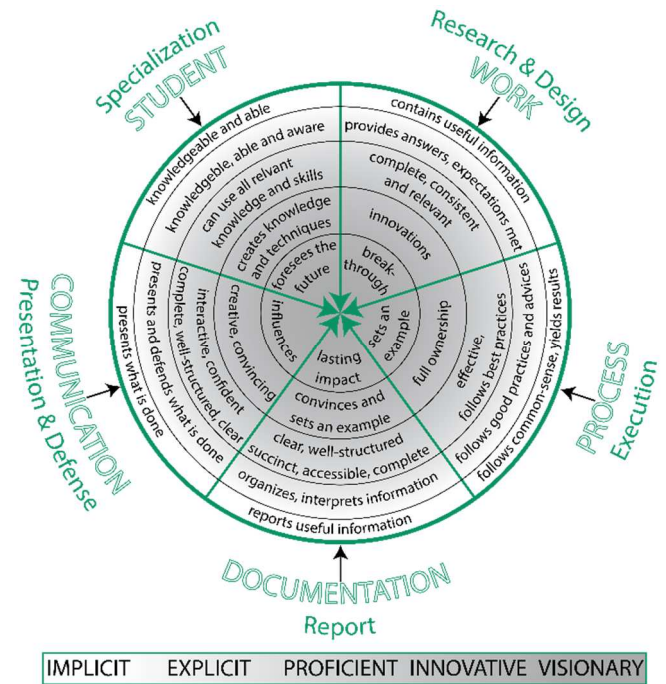
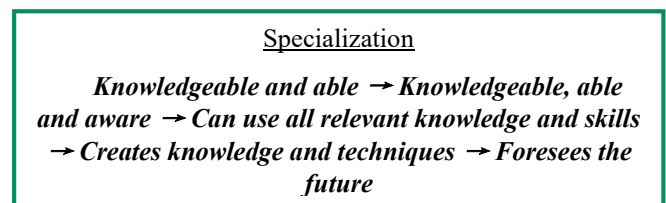


Figure 1. High-level gradients for the 5 categories of evaluation.

The scale *Implicit-Explicit-Proficient-Innovative-Visionary* is abstract but becomes concrete with respect to the category of assessment. The gradients read from the periphery of the circle, i.e., from a grade 6 to a grade 10 in the origin.

##### 1) Specialization

The category Specialization puts the student in the center and grades the performance according to the following gradient:



The Implicit Specialization level entails the following. The student knows the terminology, the specific facts and can execute algorithms in order to reach a goal. The student has the skills needed for the field. For the Explicit Specialization level, the student comprehends the challenges and can demonstrate skills and understanding of facts and ideas. The student is aware of the expectations and makes an explicit effort to meet these. To consider the Proficient Specialization level, the student has all relevant knowledge and all the skills needed to analyze and organize all the needed elements and relationships in the context of given technologies and

applications. For the Innovative Specialization level, knowledge and skills go beyond what is expected. The student proposes new theories and methods. She/he positively surprises. The student is strong in the synthesis of original solutions and in the evaluation of alternatives. The Visionary Specialization level is considered when the student shows a well-founded vision. The student shows skills which will help her/him become a leading expert in the field.

### 2) Research and Design

The category Research and Design (R&D) puts the work in the center and grades the results and deliverables according to the following gradient:

#### Research and Design

***Contains useful information → Provides answers, expectations met → Complete, consistent and relevant  
→ Innovations → Breakthrough***

The Implicit R&D level is achieved when the work is of sufficient quality, but results and contributions may be implicit. The work contains useful information and experts can extract facts, findings, results, concepts, or answers from it. For the Explicit R&D, answers are given, and expectations are met explicitly. The work features all needed aspects, e.g. comparisons, interpretations, descriptions, and statements on the main theories, concepts, techniques, and methods. The Proficient R&D level requires that the work effectively solves problems by applying relevant theories, techniques, and methods. The work analyzes all relevant literature, relationships, and results to produce evidence and support conclusions. The Innovative R&D level is obtained when there are novel results. The work goes beyond what is expected. It contains clear innovative contributions and challenges/advances the state-of-the-art. Finally, the Visionary R&D level is achieved when there are breakthroughs that will have a lasting impact, leading to potentially seminal work in a field of research.

### 3) Execution

The category Execution puts the process in the center and grades concepts like methodology, planning, attitude, which are scaled as:

#### Execution

***Follows common-sense, yields results → Follows good practices and advice → Effective, follows best practices  
→ Full ownership → Sets an example***

The Implicit Execution level requires that the project is successful. The project follows common-sense steps and produces results. The Explicit Execution level needs an explicit workflow of the student. The project flow is a result of an explicit effort. The execution follows known good practices and recommendations. For the Proficient Execution level, the execution is effective and follows known best practices. The needs of the project are understood and are reflected in the phases of the execution, leading to a smooth operation and good results. The Innovative Execution level needs the student to take full ownership of the project. The project flow is effective, pro-active, and creative. The execution suggests what to do and what better not to do. Finally, the Visionary Execution level expects that the project execution inspires and engages others. It will be given as an

example, in terms of management, results, professional network, methodology, planning, timing, cooperation, initiative, enthusiasm, etc.

### 4) Report

The category Report puts the documentation in the center and grades the actual written thesis and information that can be extracted from it. The qualities are scaled as:

#### Report

***Reports useful information → Organizes, interprets information → Clear, well-structured, succinct, accessible, complete → Convinces and sets an example  
→ Lasting impact***

For the Implicit Report level, the student report should contain information that is implicitly present in the content. Experts can extract useful information, data, concepts, methods, and results. The terminology and facts are correct. The report aligns with the established knowledge. For the Explicit Report level, the report should contain information which explicitly addresses the key points of the work. Generally, the report meets expectations in terms of depth, organization, and interpretation of the findings, arguments, and conclusions. The Proficient Report level requires the report to be clear, well structured, succinct, accessible to experts and non-experts alike, while being complete in formalization, analysis and application of knowledge to solve well-defined problems. The Innovative Report level expects that the report convincingly presents the work. The report goes beyond what is expected and can serve as an example. The report makes clear contributions and justifies these theoretically, experimentally, and against other works. Finally, the Visionary Report level expects that the report targets a broad audience and shows excellent didactic qualities communicating the innovations in an effective way. These are the graduation theses that will have a lasting impact.

### 5) Presentation and Defense

The category Presentation and Defense (P&D) puts the interactive communication in the center, according to the following gradients:

#### Presentation and Defense

***Presents what is done → Presents and defends what is done → Interactive, confident, complete, well-structured, clear → Creative, convincing → Influences***

The Implicit P&D level needs that the on-stage performance presents what has been done and contains the needed information, which is implicit in the presentation but can be extracted by the committee. The Explicit P&D level expects that the on-stage performance of the student explicitly clarifies and defends the key points of the work. Most expectations are met, e.g., non-verbal communication, message, structure, language, visuals, and connection with the audience. For the Proficient P&D level, the presentation of the work is clear, tuned to the audience, and covers all relevant aspects. The discussion and interaction with the committee show confidence. For the Innovative P&D level, the on-stage performance of the student should be convincing, original, and creative, providing new insights on the work. The presentation and defense go beyond what is expected. Finally, the Visionary P&D level requires that the on-stage performance

is influential. These are the talks and defenses that will have long-lasting impact. They will be mentioned as an example for their clarity, enthusiasm, effectiveness, material, form, etc.

### B. Aspects

Each of the five categories can be expanded, as needed, into four sub-categories that provide different aspects. We define 20 additional qualitative scales for further refining and, even more importantly, *justifying* the grade. These scales represent instantiations of the main abstract scale, too. The qualitative scales for the four aspects of the *Specialization* category are detailed in Fig. 2.

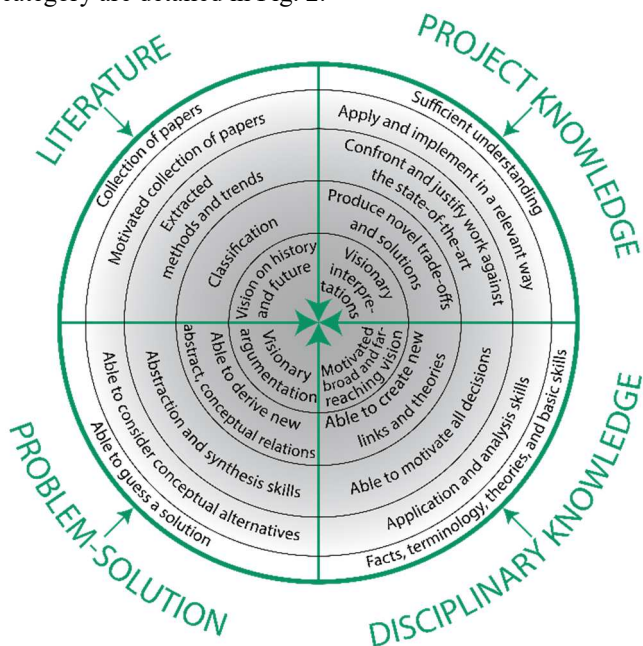


Figure 2. Gradient scales of the rubrics for Specialization

The *Research and Design* category features four aspects and each of these has again its own instantiation of the abstract qualitative scale, shown in Fig. 3 as an infographic.

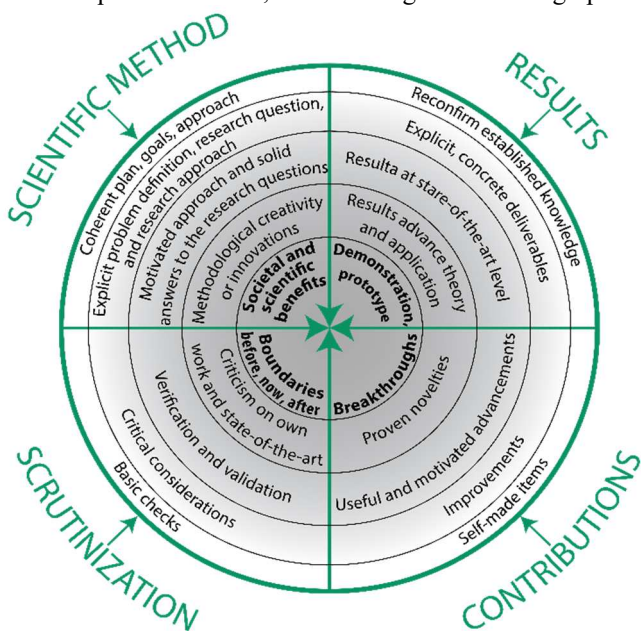


Figure 3. Gradient scales of the rubrics for Research and Design

The *Execution* category has four rubrics with their instantiations of the abstract qualitative scale. Figure 4 gives an infographic of the gradients.

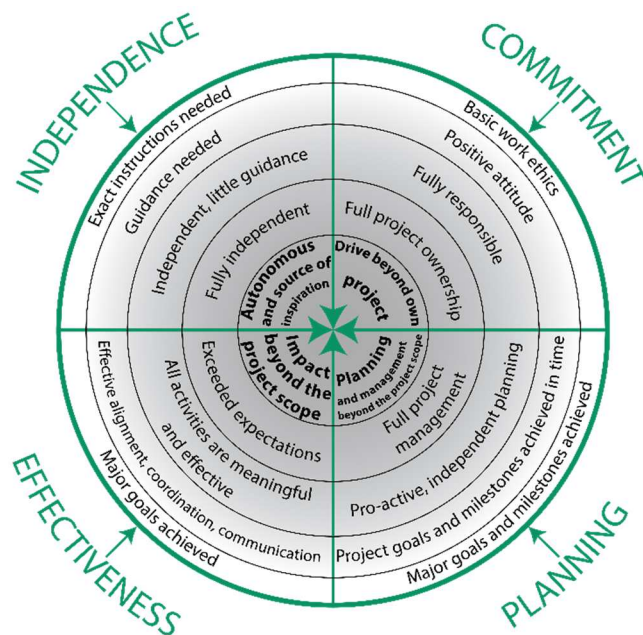


Figure 4. Gradient scales of the rubrics for Execution

Also the *Report* category has four rubrics with their instantiations of the abstract qualitative scale. Figure 5 provides an infographic of the gradients.

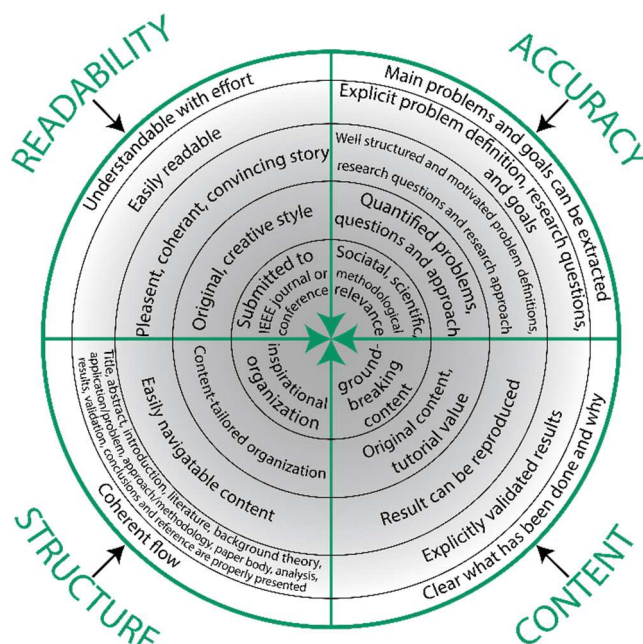


Figure 5. Gradient scales of the rubrics for Report

Finally, the *Presentation and Defense* category features four rubrics. Figure 6 shows an infographic its gradients.

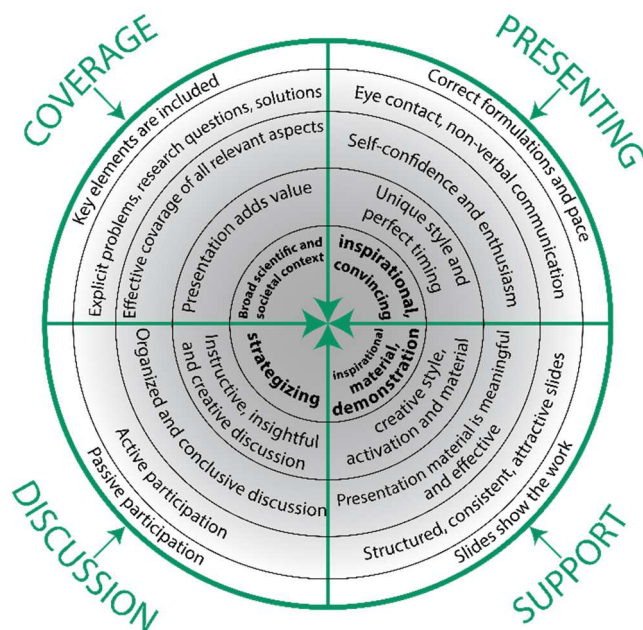


Figure 6. Gradient scales of the rubrics for Presentation and Defense

## V. OPEN ACCESS

To facilitate the access, utilization and dissemination of our guidelines, a web-app was created that contains an easy-to-use interface, provides basic explanations, and generates a further editable document of the assessment form used in our faculty. The web-app is accessible via the following link: <https://examination.master.ele.tue.nl/>.

To openly disseminate our approach, we developed a web-site, <http://eduday.wirelessresearch.nl/tiki-index.php>, that discusses our approach in detail, gives many examples, as well as further insights, and collects feedback for future improvements.

## VI. CONCLUSIONS

Fair, time-efficient, uniform, reproducible and repeatable, and conducive to learning grading requires a transparent, multi-dimensional assessment, which we approach via the proposed abstract qualitative scale:

### *Implicit-Explicit-Proficient-Innovative-Visionary.*

For the grading of our MSc graduation projects, we concretize this qualitative scale for five categories, which represent different dimensions of looking at MSc projects: a) student-centric; b) work-centric; c) process-centric; d) documentation-centric; e) communication-centric. Each of these categories has further four aspects that use their own instances of the qualitative scale. Thus, we propose in this paper 24 conceptual gradients that associate specific qualities to the assessed projects and hence lead to final grades that benefit from grade uniformity within a diverse faculty, that require limited workload from the assessors, and that support transparent communication with our students. We are currently monitoring the effects of the presented grading system. We plan to report on the results in a later publication.

## VII. REFERENCES

- [1] Winstone, N. E., & Boud, D. (2022). The need to disentangle assessment and feedback in higher education. *Studies in higher education*, 47(3), 656-667.
- [2] van Berkel, H., & Wijnen, W. (2010). Accreditation in the Netherlands: does accountability improve educational quality?. *Research in Comparative and International Education*, 5(1), 88-97.
- [3] Hammer, S. (2007). Demonstrating quality outcomes in learning and teaching: Examining 'best practice' in the use of criterion-referenced assessment. *International Journal of Pedagogies and Learning*, 3(1), 50-58.
- [4] Pathirage, C., Haigh, R., Amaratunga, D., & Baldry, D. (2007). Enhancing the quality and consistency of undergraduate dissertation assessment: A case study. *Quality Assurance in Education*.
- [5] Walser, T. M., & Trevisan, M. S. (2016). Evaluability assessment thesis and dissertation studies in graduate professional degree programs: review and recommendations. *American Journal of Evaluation*, 37(1), 118-138.
- [6] Grainger, P., Christie, M., Thomas, G., Dole, S., Heck, D., Marshman, M., & Carey, M. (2017). Improving the quality of assessment by using a community of practice to explore the optimal construction of assessment rubrics. *Reflective Practice*, 18(3), 410-422.
- [7] Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159-179.
- [8] Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & evaluation in higher education*, 35(4), 435-448.
- [9] Christie, M. F., Grainger, P., Dahlgren, R., Call, K., Heck, D., & Simon, S. (2015). Improving the quality of assessment grading tools in master of education courses: A comparative case study in the scholarship of teaching and learning. *Journal of the Scholarship of Teaching and Learning*, 15(5), 22-35.
- [10] Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational research review*, 9, 129-144.
- [11] Bloom, B. S., 1956. Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc.
- [12] Newmann, F. M. *Authentic achievement: Restructuring schools for intellectual quality*. Jossey-Bass, 1996.
- [13] Close, Daryl. Fair grades. *Teaching Philosophy* 32.4 (2009): 361-398.
- [14] Sullivan, T. A. Greedy institutions, overwork, and work-life balance. *Sociological Inquiry* 84.1 (2014): 1-15.
- [15] Rao, A. S. (2020). Towards Ensuring a Greater Global Uniformity and Standards in Higher Education. *International Journal of Education, Culture and Society*, 5(6), 141.
- [16] Wihlborg, M., & Teelken, C. (2014). Striving for uniformity, hoping for innovation and diversification: A critical review concerning the Bologna process—Providing an overview and reflecting on the criticism. *Policy Futures in Education*, 12(8), 1084-1100.
- [17] Boud, D., Freeman, M., James, R., Joughin, G., Sadler, R., Dochy, F., & Fitzgerald, T. (2010). Student assessment for learning in and after courses. *Sydney, NSW: Australian Learning and Teaching Council*, 1-31.
- [18] Clementsa, M. D., & Cord, B. A. (2013). Assessment guiding learning: developing graduate qualities in an experiential learning programme. *Assessment & Evaluation in Higher Education*, 38(1), 114-124.