METU JFA 2026/1
(43:1) *277-300*

DOI: 10.4305/METU.JFA.2026.1.10

DEVELOPING A RUBRIC TO ASSESS STUDENT WORKS IN FOUNDATION DESIGN COURSE

Dalsu ÖZGEN KOÇYILDIRIM*, Damla TONUK**

Received: 12.05.2025; **Final Text:** 20.04.2026

Keywords: Design education; foundation design; rubric based assessment; rubric for design studio; studio pedagogy.

* Department of Industrial Design, Middle East Technical University, Ankara, TÜRKİYE.

** *Corresponding Author;* Department of Industrial Design, Middle East Technical University, Ankara, TÜRKİYE.

INTRODUCTION

Assessment in foundation design education - the first year studio course where the basics of design elements, principles and processes are laid out - is often a complex and chaotic process bound in tradition that does not usually comply with current understandings of assessment for learning. Moreover, aligning education constructively requires an open assessment system (Gelmez and Arkan, 2021). Due to design's uncertain quality (Schön, 1983), assessment is often not clear for outsiders and students, and it even appears to be subjective, which is frequently problematized by scholars (Schelling et al., 2012; Bartholomew et al., 2019; Sudhindra and Blessing, 2021; Taştan and Er, 2025). With this study, we aim to improve assessment in the foundation studio courses by bringing much-needed systematization and theoretical grounding so that learning goals are better communicated with students and students are more active in planning their own projects.

Assessment of student works in foundation studios traditionally follows "the connoisseurship model of assessment" (Parkes, 2010; Smith, 2013) where expert tutors collectively discuss student works one by one by comparing them and discussing qualities and aspects of each work, such as whether it answers the given problem or has a good compositional structure. This is a tedious process where consensus is not always achieved, or may require long discussions, which risk shifting the focus away from the learning goals. Consequently, the assessment appears subjective and based on individual valuation as the criteria for assessment are not visible nor conveyed to students (Kasap and Kaptan, 2022). Indeed, as foundation studio tutors, we receive feedback about the transparency and clarity of assessment criteria, and are therefore concerned that our assessment methods do not contribute to student learning as underlined in current understandings of "assessment for learning" or "formative assessment" approaches (Bennett, 2011).

As Kasap and Kaptan (2022) discuss in their study, even a subjective matter such as design may be objectified by expertly defining assessment criteria. Therefore, we believe that a rubric that lays out and quantifies the criteria through which the successful attainment of the learning goals may be judged, can offer us a way to systematize and objectify our assessment. In the literature, there are examples of studies for assessment and rubric development in the education of art disciplines, architecture, interior architecture, fashion design, industrial design, liberal arts, information literacy, medicine, nursing, management, dentistry, food technology, teacher education and film technology (Reddy and Andrade, 2010; Nuhoğlu et al., 2017; Kasap and Kaptan, 2022; Yuan et al., 2025) but not specifically foundation design education. Therefore, we undertook a study to remedy that lack; the process took three years, which consisted of five iterative stages.

This study contributes to Industrial Design education and design education literature by arguing for the validity of systematized assessment in foundation studio courses and explaining how assessment criteria may be understood and defined in that specific context. We provide the rubric we developed but we also believe that the account of our process and insights will be useful to other educators in developing their own rubrics since the teaching methods and intended learning outcomes of foundation studios are similar even though the specific exercises and tasks may differ from one course to the next. Furthermore, we hope that any other discipline that needs to assess visual outputs of any kind can benefit from this study, as visual communication is relevant to many non-design disciplines.

In what follows, we explain the specifics of design studio education, and more specifically foundation design studio. Then, we set the theoretical framework of how assessment is to be understood in this setting and explain our own rubric development process. Finally, the findings of inter-rater reliability analysis and our field tests will be presented to discuss the use of rubrics in foundation studio, along with its limitations and future applications.

STUDIO PEDAGOGY AND FOUNDATION DESIGN STUDIO

Design education is commonly conducted in studios with a "signature pedagogy" (Schulman, 2005; Crowther, 2013) that originates from the Bauhaus school, where students learn by doing and experimenting (Gür and Yüncü 2010; Acar et al., 2019), i.e., working on actual design problems similar to the ones students will face in their professional life. Students proceed in these projects by getting feedback - called critiques - that do not aim to direct them to one true answer but rather guide them in finding their own individual solution. As a result of this emphasis on the individuality of the learning process, studio education is also defined as a "flexible pedagogy" (Biggs, 1999; Rohse and Anderson, 2006). Boling and Schwier (2016, 13) explain that the studio is "activity-centric" but unlike in traditional education "none of the activities is designed to teach content explicitly or directly". This educational approach aligns perfectly with constructivist learning environments that are defined as student centered, explorative, encouraging of new ways of thinking, collaborative, focused on hands-on problem-solving activities relevant to real life, and where the emphasis on the process rather than the final result leads to concurrent learning and assessment (Arkün Kocadere and Özgen, 2012).

Within this general context of design education, the first year studio course, in some geographies referred to as the Basic Design Course or more broadly as Foundation Design Course makes the first year of the programme as an initiation class to visual education, where visual "composition language and rules" - in other words, design elements and principles - are learned by doing (Acar et al. 2019, 76). Though the methods and exercises might vary from one institution to the next, the main educational approach where "instructors set up circumstances that encourage habits of thought and action, promote reflection" (Boling and Schwier, 2016, 13) remains similar. Through these, students gradually become more conscious of visual outputs and gain the ability to compose and structure these purposefully.

The specific foundation studio where we based our study is conducted at the Industrial Design Department of Middle East Technical University and is a module that spans the full first year of the undergraduate education, conducted over 12 hours a week, during a total of 28 weeks, where students work on a variety of projects that sequentially build up on complexity. The course is structured as four consecutive periods: (Q1) it starts with 2D explorations of basic design principles, (Q2) followed by the application of these principles onto relief designs as an initial step towards a (Q3) full on exploration of three dimensional forms and volumes, and (Q4) ending with a combination of 2D, relief and 3D compositions where concepts such as function and user start to be incorporated.

To begin the systemization of assessment, we started with Q1 - the initial seven weeks where we aim to teach students the visual elements and principles of design and composition: Unity, repetition, rhythm, dominancy, contrast, harmony, figure-ground, transparency, empty space, are commonly referred to as principles that organise a visual space (Demirkan and Afacan, 2011; Acar et al., 2019). The 2D assignments executed in that period are short projects, usually executed with basic geometric shapes cut out of cartons, produced in two to three hours. As an example, we ask students to make a 2D organisation in a given area that emphasises direction using basic geometric elements, and we expect the resulting work to both showcase this principle and to configure elements in relation to each other and to the area.

Exercising with these compositional elements and principles is an integral component of developing creativity (Demirkan and Afacan, 2012) hence they are fundamental for design education. The result is a composition work, not a design product, but teaches visual sensitivity, basic principles of visual organisation, craftsmanship, attention to detail, and willingness to explore, which are observed in the physical output. Assessment of whether the student achieved these is usually done by two or more tutors discussing their point of view, based on their expertise and sometimes comparing the whole group or often comparing works as needed to define their ranking. We have personally observed that this process does not always end in consensus, and sometimes, one tutor's emphasis on one aspect shifts the assessment focus, or gives too much importance to one aspect at the cost of disregarding merits of other aspects. As a result, we felt the need to clarify first what exactly we look for to judge whether a student learned what was intended - the specific criterion of assessment - followed by how to coordinate multiple criteria fairly and consistently when assessing.

HOW TO ASSESS FOUNDATION DESIGN STUDIO WORKS?

Studies on design education indicate that assessment in the studio course needs to be compatible with the intended learning outcomes (Ellermers et al., 2008; Giloi and du Toit, 2013), just like in any outcome-based teaching and learning (OBTL) (Biggs and Tang, 2011), which relies on the coordination of the intended learning outcomes with the educational strategy, the assessment strategy and the assessment criteria (**Figure 1**).

Adopting the "criterion referenced" assessment of OBTL enables educators to observe how well individual students achieve the intended outcome and prevents grading comparatively (Biggs and Tang, 2011, 11).

Even though there are contradictory opinions on whether the intended learning outcomes and corresponding assessment criteria can or should be clearly and fully determined for the studio courses, due to the uncertain nature of design education (De La Harpe et al., 2009; Seery et al., 2011; Davies, 2012; Orr and Bloxham, 2012), a significant number of studies show that a clearer definition of the assessment methodology and criteria is very helpful in achieving consistency between the educational team as well as any guest assessors (Davies, 2003; Rust et al., 2003; Grobler, 2006; Webster, 2007; Çıkış and Çil, 2009; De La Harpe et al., 2009; Uzunoğlu and Uzunoğlu, 2011). To comply with our aim to improve the assessment in our foundation design course, we need to first investigate how to define the assessment criteria to match the intended learning outcomes of our course, and then decide on the assessment strategy - the method by which to measure it. The focus of our study being the Q1 period as stated above, we needed assessment criteria specific to the learning outcomes of Q1: *composing* 2D abstract visual works with basic design elements and principles, to *creatively solve* basic design problems.

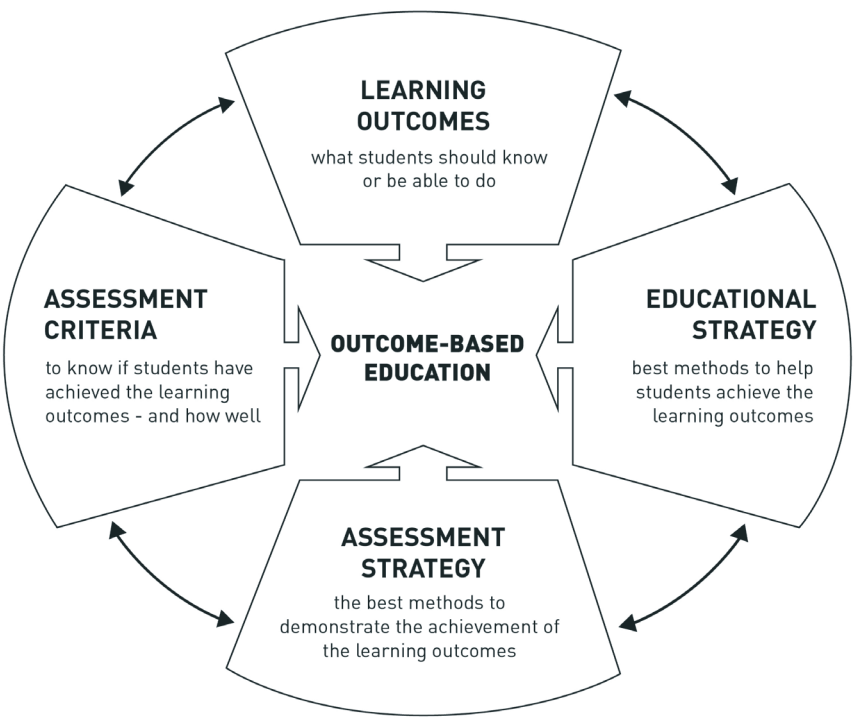

Figure 1. The relationship of learning outcomes, educational strategies, assessment strategies and assessment criteria in OBTL, (adapted from Bingham, 2002).

Assessment Criteria

A review of scholarly accounts of how assessment criteria for design courses have been defined and used directed our quest in improving assessment in our foundation studio course. The literature revealed varied approaches to defining assessment criteria: drawing criteria from literature if possible (Demirkan and Afacan, 2011; Ketizmen and Güçyeter, 2024), interviewing and discussing with experts (Webster, 2007; Nuhoğlu et al., 2012; Hasdoğan, 2012; Brookhart, 2013; Onwuagboke and Singh, 2016), observing experts during evaluation (Ketizmen and Keleş, 2025), statistically analysing job specifications (Yuan et al., 2025), and discussing with students (Karahanoğlu, 2022).

Studies have been conducted on how to define assessment criteria of creativity for first year interior design students (Demirkan and Afacan, 2011), first year architecture students (Ketizmen and Güçyeter, 2024), design thinking (Aflotony, 2018), finished products (Schelling, 2012) and design portfolios (Tastan and Er, 2025). For example, Hasdoğan (2012) provides a detailed account of developing assessment criteria for a government initiated national design awards scheme, where they use two rounds of expert interviews. Similarly, Schelling et al. (2012) devised assessment criteria to evaluate graduation projects that comprised evaluating competency of the final outcome in more dimensions than just the product, but also a report, presentation, and real-life setting. However, there are no accounts of an attempt at defining assessment criteria for foundation design nor at the systematization of assessment in general for the first year of design education, in whichever name or format it might be. The above-mentioned accounts, when they mention elements of visual design principles either take it as a too general concept, or take it solely in terms of creative or novel solutions.

Nonetheless, two studies stand out in terms of the similarity of the outcome studied with our studio course. Demirkan and Afacan (2011) focus on 2D artefacts in foundation studio course work somewhat similar to the 2D work produced in our studio. However, their focus is on creativity as they aim to define how affective characteristics of creativity correspond to design elements and principles. However, our aim in Q1 is to learn and explore these elements and principles, not to assess how these principles relate to creativity. Onwuagboke and Singh (2016) focus on graphic design artefacts and they derive assessment criteria through expert discussions. Although they are focusing on 2D works, their criteria have additional emphasis on communication of ideas, and use of technology. Their criteria do not detail elements of the visual organisation so these criteria could not be employed in our studio. As such, these examples could only partially provide us with assessment criteria that are relevant to our context and therefore, more exploration was needed to obtain definitions that fully encompass the intended learning outcomes of our course.

The literature also provided us with additional suggestions on this process. Webster (2007) proposes that, to conduct an expert discussion, teachers compile works from previous years and by going over this portfolio, they discuss how they grade individual projects, in order to collect an initial list of keywords as evaluation criteria. Nuhoğlu et al. (2012) further explain that "the criteria determination cycle consists of the stages of reflecting, listing, grouping and labelling". Brookhart (2013, 33) advises leaving "descriptions open to professional judgment" and states that "making some inferences—is better than locking things down with overly rigid

descriptions". This arguably also leaves room for expert evaluation and prevents the assessment from becoming too mechanical.

Assessment Strategy: The Rubric

In accordance with OBTL, what is expected of students needs to be explicitly stated, defined, ranked and the SOLO taxonomy, which maps out the successive levels of a learner's cognitive development, is a useful guideline to organize and understand these expectations. The SOLO taxonomy, widely adopted in both educational sciences and design research (Leung, 2000; Roy et al., 2022), proposes that learning and understanding grows in a hierarchical fashion, building up with the acquisition of increasingly complex knowledge, skills and ways of thinking. In this hierarchy, "where each partial construction becomes the foundation on which further learning is built" (Biggs and Tang, 2011, 90), the progress is expressed in five succeeding steps that range from low to high cognitive levels. **Figure 2**, adapted from the studies of Biggs (2013) and Arı (2013), shows the details of these steps: the learner gradually goes from the initial state of having no knowledge to being able to synthesize and apply the acquired knowledge in different ways and other contexts. In the same figure, a hierarchy of verbs that explain what "students need to be able to do to indicate achievement at the level in question" (Biggs and Tang, 2011, 90) is also given. The target verbs given above for the intended learning outcomes of the Q1 period can be seen at the highest level: to compose and creatively solve.

This taxonomy lends itself to the creation of a rubric, a grading tool commonly defined as "a document that articulates the expectations for an assignment by listing the criteria or what counts, and describing levels of quality from excellent to poor" (Reddy and Andrade, 2010, 435). Rubrics lay down assessment criteria in a hierarchical fashion, where the final levels of understanding reflect the course's intended learning outcomes in terms of their structural quality" (Biggs and Tang, 2011, 81). "Rubrics and the use of rubrics is currently one of the 'hottest' research topics in education and

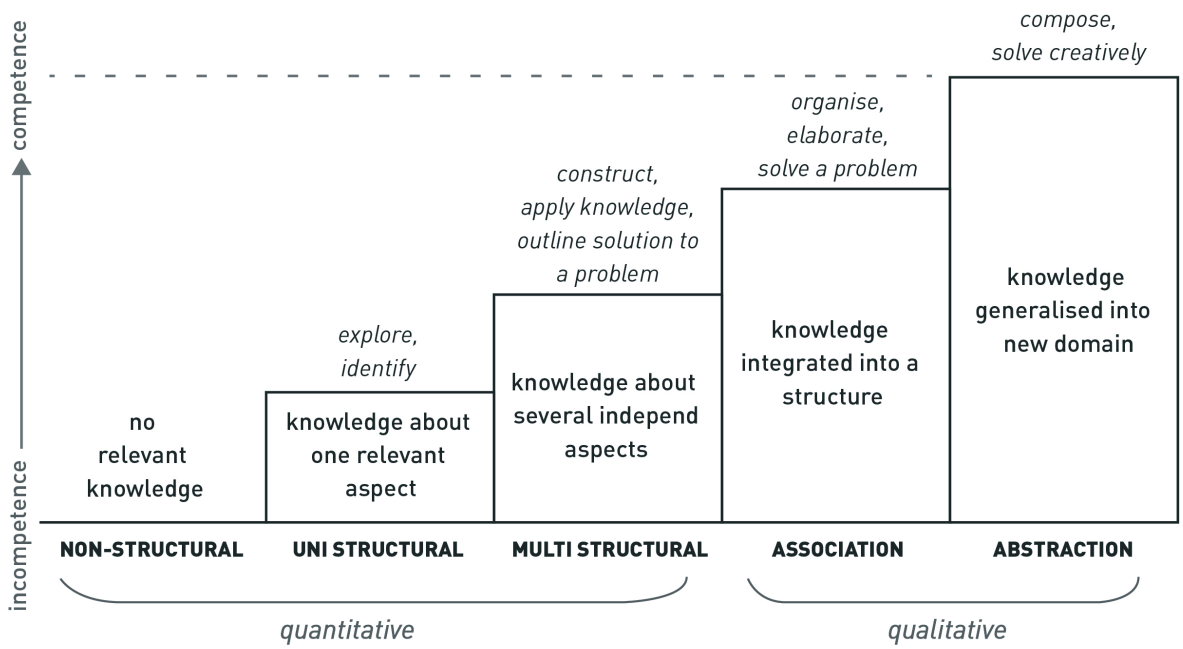

Figure 2. Cognitive development of the learner, adapted from the SOLO taxonomy (Biggs and Tang, 2011), target verbs in Italic are defined for our Q1 period.

education psychology" claim Panedero and Jonsson (2020, 1) based on their evaluation of number and venue of publications of recent years on this topic. However, this does not imply that all the publications support the use of rubrics, especially for design education.

Three main sub-themes were identified as points of concern in these works: Firstly, scholars are worried that "the steering effect of assessment criteria" will result in students only trying to match the criteria in a superficial manner, rather than internalizing knowledge (Bearman and Ajjawi, 2018; Kohn, 2006, as cited in Panedero and Jonsson, 2020, 25). Secondly, laying out criteria in a transparent manner is believed to remove learning challenges, resulting in "extensive coaching" (Torrance, 2007, as cited in Panedero and Jonsson, 2020, 13) and therefore, eliminating the learning that happens on the way to discovering what the criteria are. Finally, this lack of self-discovery leads to the "transfer of epistemic beliefs" which gives students a "sense that knowledge is fixed and stable" (Bearman and Ajjawi, 2018, as cited in Panedero and Jonsson, 2020, 26), obstructing the development of inquisitiveness, creativity and the ability to work with uncertainty.

However, there are well-grounded studies with empirical research that demonstrate the benefits of rubrics for assessment. Panadero and Jonsson (2014) contend that if a rubric is designed and implemented in line with the intended outcomes, it can have positive effects on students' learning, academic performance, and self-regulation; they "can teach as well as evaluate" (Andrade and Du, 2005, 1). A rubric has the potential of achieving all these because it clearly states what is expected of students according to the assessment for learning approach of student-centered education, and as a result, students will better understand the assessment they received, may self-assess more successfully and become more in tune with both their instructors and the course environment (Bohemia et al., 2009; Webster, 2007; Rust, 2002). It was revealed that students can organise their work and plan their approach better, as they can check their work on required criteria and so are able to produce more quality work and get better grades (Reddy and Andrade, 2010), and not get as anxious about their performances (Andrade and Du, 2005). Indeed, a study in a post-graduate level course comparing the grades of two groups, one of which was presented with a rubric at the beginning of their project, clearly showed that the group with the rubric got higher grades (Petkov and Petkova, 2006). As for the instructors, they claimed that rubrics eased and shortened their tasks (Reddy and Andrade, 2010), helped them to grade "more consistently, reliably and efficiently" (Campbell, 2005) and ensured consistency between the educational team as well as any guest assessors (Uzunoğlu and Uzunoğlu, 2011; De La Harpe et al., 2009; Çıkış and Çil, 2009; Webster, 2007; Grobler, 2006; Davies, 2003; Rust et al., 2003). As Giloi and du Toit (2013) point out, such consistency allows educators to assess the efficiency and effectiveness of the learning they provide.

In the light of the benefits of rubrics outlined above, we are in line with Brookhart's inference: "[Some scholars may] argue that rubrics constrain creativity and metacognitive development. I disagree. Rather, bad rubrics constrain creativity and metacognitive development" (2013, 27). Accordingly, we hope that a well-designed rubric will still allow space for the unexpected and the creative while getting rid of the black box myth of an expert evaluator.

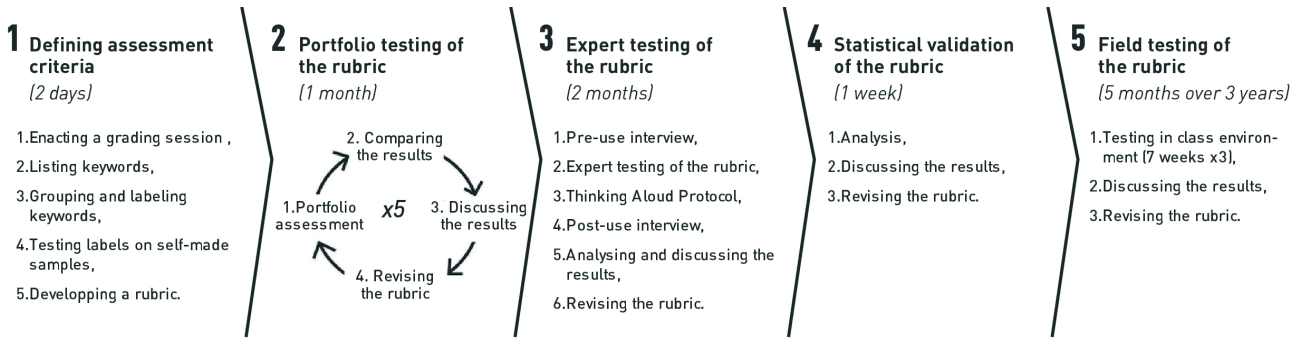

Figure 3. The five stages of our rubric development.

METHOD: THE PROCESS OF RUBRIC DEVELOPMENT

Schön (1983) has famously defined design as a reflective practice that deals with uncertainty: researching, discovering and experimenting in an iterative manner within fluid and blurry boundaries are of its essence. Accordingly, Gasaymeh (2011) states that a rubric for such an uncertain topic needs to be general instead of task-specific, to allow for creativity, for assessment of consecutive assignments and different contexts, as well as flexible, so that the rubric is not a fixed checklist but an ongoing process, adapting to the needs and changes of the instruction.

The development process consisted of five main stages (**Figure 3**), with the second stage itself consisting of five iterations. We started our development process after acquiring ethical approval.

Defining Assessment Criteria

Following the suggestions found in the literature as explained above, we started the first stage of defining the assessment criteria by conducting an expert discussion on a portfolio of 100 randomly selected student works from previous years, produced during the Q1 period. As trained designers who have been teaching the same foundation studio course for more than 10 years, we acted as the experts: we enacted a grading session and then discussed the aspects we usually consider while assessing a work. We gathered an initial list of keywords out of our discussions, such as use of space, relationship between elements, symmetry or asymmetry. We soon realized that some of these keywords were referring to a perceived problem, some defined a clear category and others indicated different levels of quality within the same category. As Nuhoğlu et al. (2012) suggested, we went through an iterative process of reflecting, listing anew, grouping-labelling and reflecting once again to better conceptualize the categories and bracket the terminology.

We paid specific attention to the idea that the rubric had to allow creativity, different or unexpected solutions to better comply with the nature of design education (Brookhart, 2013, 33). As a result, we endeavoured to label and define evaluation criteria as widely as possible and to suggest multiple options.

To distinguish the different competence levels within each criterion, we developed a method where the list of keywords gathered previously were used as variables in a sequence of matrices. Sample 2D compositions were used to visualize the interaction of these variables so that we may discuss which combinations rate higher in competence. For example, while one composition displayed an asymmetric organization of varied elements with a straight axis (**Table 1a**), a second one showed the same elements with the

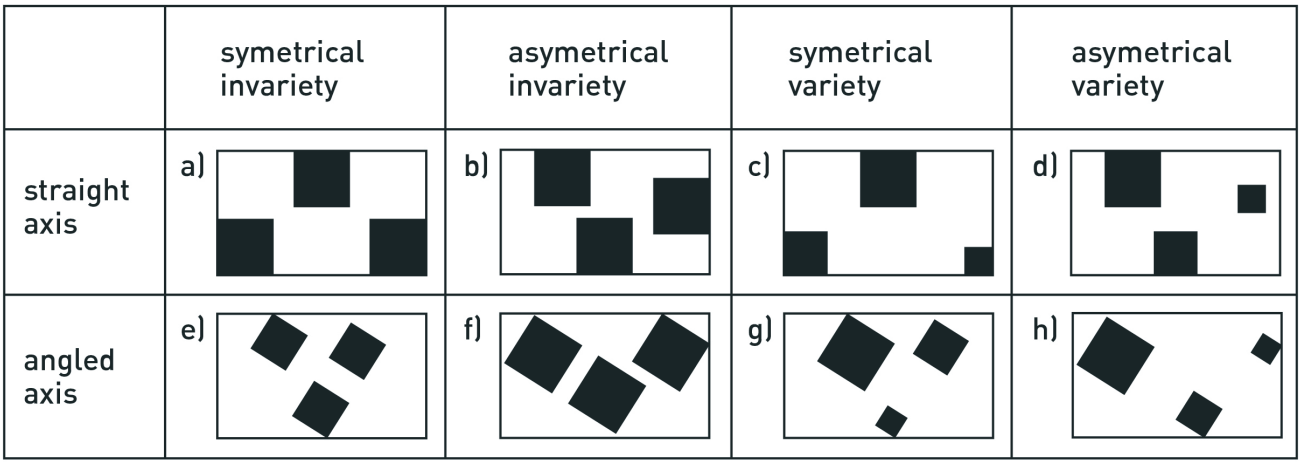

Table 1. Matrix of interactions for three variables: Symmetry, Variety, Axis.

same axis but in an asymmetrical order (**Table 1b**) and a third one had a single type of unvaried elements in an asymmetrical layout with an angled axis (**Table 1e**). The sample designs were compared and rated as higher or lower in terms of levels of competence in accordance with our intended learning outcomes: to achieve the ability to compose, we expect students to handle a variety of basic design elements together in a multi structural way, which, in the example given here, would stand for asymmetrical variety with an angled axis.

Furthermore, we attempted to assign grading points for the different items on the rubric, so that all assessment criteria would have equal weight as well as each level within would have equally distributed points. The points were adjusted to grant a pass grade (55 over 100 in our university) to projects that wererated in the second competence level while below that level, projects are deemed as failed.

By the end of this phase, we obtained a draft rubric with five assessment criteria, and five levels of competence (see Appendix A).

Portfolio Testing of the Rubric

The second stage of our process consisted of testing the draft rubric on the selected portfolio of 100 student works, comparing the results, discussing its effectiveness, revising to overcome the issues and testing once again. Over a month we did five iterations to prepare a more refined version that could be tested on a wider scale by different tutors. The purpose of these iterations was to construct a more balanced rubric, where all criteria have equal importance, and are divided into an equal number of levels with the same grade value. Each criterion needed to be structured to "describe a continuum of levels of performance quality. These levels [had to] be distinguishable [… and] able to describe what is different from one level to the next" (Brookhart, 2013, 27). As a result, new labels were introduced and new groupings were made. For example, the concept of asymmetry in a composition had figured quite frequently in our discussions which led us to define it as a criterion of its own. However, as we tested the draft rubric, it started to fit better as a development level of a wider criterion dealing with the structure of the composition. We also became aware of other aspects beyond our initial keywords that underlies our assessment, such as the complexity of the relationship between varied elements. As we refined the competence levels of each criterion, we also aimed to include the course's intended learning outcomes for the Q1 period as the most

advanced level. Therefore, by the end of this stage the final column of the rubric more closely expressed what students should strive to produce: creative solutions to a given problem, by organizing visual elements using basic design principles.

At this stage, our initial attempt at designing a rubric for foundation design was presented at a national design research conference (Özgen Koçyıldırım and Tonuk, 2020). We received positive feedback from fellow foundation studio lecturers, and requests to use our tentative rubric for assessment in their own classes.

Expert Testing

The tentative rubric (Appendix B) was tested with 11 experts, whose expertise ranged from maximum 30 years to minimum 5 years of teaching on this topic. They were, at the time of the study, actively teaching foundation design to industrial design students in six different universities in two different cities. The rubric was mailed in advance so that the experts could familiarize themselves with the rubric. In the trial sessions, they were asked to assess 20 student works selected from previous terms by using the rubric. During the assessment the experts were asked to assess the first 10 works on their own, and for the last 10 works, a think aloud protocol was applied: they were asked to verbalize their thinking process as they use the rubric, so that we could observe how they understood the criteria and interpreted the rubric. At the end of their assessment session, we asked them to comment on their experience and the design of the rubric, and whether they had additional suggestions. In accordance with experts' comments, the rubric was revised by refining the wording and phrasing of the criteria and the competence levels as some explanations were either too vague or too detailed or too restrictive. The second criterion was re-worded as the Organisation instead of the Overall Structure and the third criterion was redefined as the Use of Acquired Knowledge and [Design] Principles as suggested by the experts.

Statistical Validation

Before moving on to the next phase, we checked the statistical consistency of the rubric: we performed an inter-rater reliability analysis using the grades submitted by the experts during the testing. IBM's SPSS was used to calculate Cronbach's alpha, a measure of the internal consistency for assessment tools, especially suited for cases with multiple assessors (Stemler, 2004). This method is advantageous because it yields an overall estimate of consistency which makes for an easier and quicker evaluation of our rubric's success, also because it does not require the assessors to be in exact agreement on how to use the rubric and therefore does not need any preliminary training of the experts. In order to use Cronbach's alpha, the rating scales need to be continuous, which they were, since the experts could and did give the exact rate they wanted within the pre-established levels of competence.

Field Testing

The Revised Rubric (Appendix C) was tested during the 2020-2021, 2021-2022, 2022-2023 and 2023-2024 academic years' fall semesters in the foundation studio of the Industrial Design department at Middle East Technical University. A total of 394 first year students (94 in 2020, 100 in 2021, 100 in 2022 and 100 in 2023) attended the course, and the rubric was used to assess approximately 2500 individual 2D works produced during

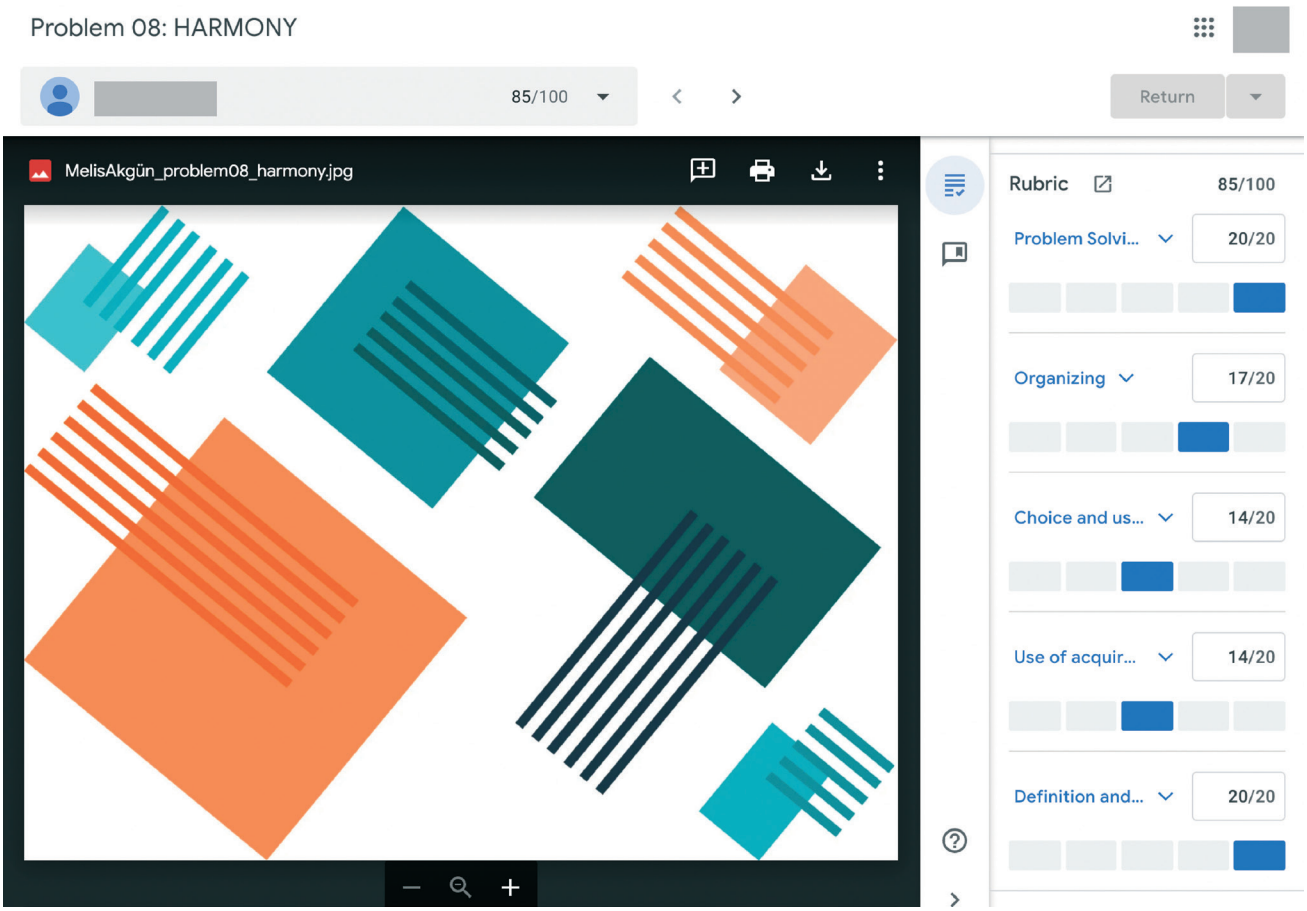

Figure 4. Google Classroom interface with the rubric used for grading.

the Q1 period. Both the submission of works and their assessment were processed through Google Classroom. This digital platform was chosen because its interface allowed for the visual observation of individual works alongside its customisable built-in rubric service where the criteria definitions and levels can be presented in a clear and understandable way (**Figure 4**). The Revised Rubric was used by seven members of the studio teaching team, including instructors and research assistants. The users were questioned periodically through the process, to gather their opinions on using the rubric for assessment.

The large amount of work that was assessed, the many weeks long duration of the testing, as well as the different perspectives of various assessors provided us with enough information and feedback on the effective or the problematic aspects of our Revised Rubric.

FINDINGS AND DISCUSSION

Inter-rater Reliability

Of the 11 experts that participated in the study, only 9 rated every item presented to them. Since the chosen method requires every item to be evaluated, the incomplete inputs of two experts were disregarded. It should be noted that according to Jonsson and Svingby (2007), who reviewed 75 papers about scoring rubrics, even two raters are often "enough to produce acceptable levels of interrater agreement". The analyzed data set consisted of the overall grades given by 9 experts to 20 student projects as

	Overall Reliability of the rubric	Reliability of Criterion 1 "Problem Solving"	Reliability of Criterion 2 "Overall Structure"	Reliability of Criterion 3 "Choice and Use of Elements"	Reliability of Criterion 4 "Use of Acquired Knowledge"	Reliability of Criterion 5 "Definition and Use of Area"
Cronbach's Alpha	**.802**	.576	.870	.691	.741	.771

Table 2. Reliability statistics of the tentative rubric.

well as the scaled scoring of the five separate evaluation criteria. In general, coefficients above 0.6 indicate a satisfactory reliability, above 0.7 is deemed as good and when they exceed 0.75, reliability may be judged as high or robust (Taber, 2017). The consistency estimates of the rubric overall as well as each criterion is presented in **Table 2**.

The Cronbach's Alpha for our rubric overall indicates a very high reliability with a coefficient of 0.802. Criterion 1, which stands for "problem solving" figures as the most problematic criterion, with the lowest coefficient. Indeed, experts had indicated certain issues with the vocabulary of its competence levels during the interviews. As a result, the term diagrammatic was replaced with conceptual, and the definitions of the criterion's level of competence were revised. The reliability of all the other criteria were found to be good.

Analysis of Expert Opinions

The comments and suggestions of experts were used to refine the final structure and vocabulary of the rubric that was further tested in the field. The general opinion about the rubric of both experts and field-testing tutors was very positive. Overall, the assessors were able to evaluate based on the provided criteria and did not feel that anything was left out. They found it useful in terms of both doing justice to student work but also in explaining to students the reasons behind the grades they receive. Mostly, they found it more efficient than the traditional methods of assessment applied in studio education: it was judged to be easier and less time consuming since it eliminated the need to compare student works among each other, and equalized the assessment approach of different instructors. In addition, students did not have any questions or need for extra explanations in the terms we used in the rubric.

The most successful aspects of the rubric were listed as:

- easier grading,
- quicker grading,
- more consistent grading,
- less need to compare between student works,
- reminding the learning outcomes,
- more detailed inspection of the work.

The points of concern were:

- lack of assessment for craft quality,
- debatable hierarchy of levels within criteria,
- assignment specific suitability of levels within criteria

Regarding points of concern, craft quality is very important in foundation year studio, as it is key to designerly understanding. However, we felt that craft cannot be taken as a whole criterion that makes up 20 points, with 5 levels, so it is better to evaluate crafts quality separately.

7 out 11 experts expressed concerns regarding the hierarchy levels and distribution or the weights among these. As we delved deeper, we realised that these concerns resulted from several aspects: First, because they were not used to using rubrics the experts had trouble in dividing levels into equal weights. Second, experts lacked the context of which week the assignment was submitted so what level to expect. Third, the levels were specific to the 2D works in Q1 period. Relatedly levels are different in the first weeks of the term and at the end of the 7 weeks, but we feel that the criteria are nevertheless applicable to assessing 2D compositions of any level in general. Furthermore, during the post-use interviews, it became clear that the problem was mainly with the wording of the rubric rather than with the way we divided and structured the levels. The experts' suggestions were very helpful in fixing these lexical problems; we changed the wording to make it more open to allow for different possible expressions and understandings of the criteria.

When discussing each criterion one by one, the experts voiced concerns primarily with criterion one (Problem Solving*)* and criterion two (Choice and Use of Elements), which reflected the results of the inter-rater reliability analysis.

Problem Solving: 8 out 11 experts found the definition of this criterion to be too wide. The first objection was about the 20/100 weight given to this criterion, since some experts considered the solving of the problem as the main task in projects, requiring at least 50% of the overall grade. After discussions, we decided that in our own specific approach, we take this criterion to be of equal value with the other criteria. Furthermore, in our case a specific design principle, such as direction, balance or dominancy, is given as the main problem in the brief, whereas in some courses conducted in other institutions, the design principles are studied within different briefs or projects. Possibly, the expression Problem Solving expressed more concerns regarding a design problem rather than visual solution within the context of basic design, so a slight correction as Problem Solution might be a better expression.

Choice and Use of Elements: Experts found it difficult to consider this separately from the previous criterion in the rubric, Overall Structure. Their comments suggested that the detailed explanations given for this criterion's competence levels, with lots of and/or attributes, worked against an easy understanding and increased their confusion. Therefore we updated the wording for this criterion's levels and started to use shorter sentences with a cleaner syntax.

The three other criteria of the rubric - Overall Structure, Acquired Knowledge /Designerly Attention, Definition and Use of Area - which had high results in the inter-rater reliability analysis received very few comments, dealing mostly with their syntax. For example, the use of the word asymmetry was found to be too restrictive or directive, and dynamism was suggested instead, replacing a structural definition with a more abstract concept.

To further improve their communication, we revised their wording as well in line with the changes made to criterion three.

The Evolution of the Applied Rubric

Through the four years it was used in the classroom, the rubric slowly evolved over time: following the studio team's satisfaction and

dissatisfaction with its use, the wording of its criteria and competence slightly changed (**Figure 5**).

The longer it was used, the less important the details of the competence levels appeared to be, as the instructors became familiar with the criteria and started to perceive them as five consecutive and cumulative levels, without focusing too much on their specific wording.

Through the four semesters of field testing, the studio team observed a few major contributions of the rubric:

- The team could more comfortably divide the large amount of assignments to grade among themselves to lessen their individual workload, without any concern of fairness and consistency.
- There was a significant decrease in the number of students questioning their grade, how the assignments were evaluated, indicating an improvement in the communication of learning outcomes and educational methods. This result was clearly due to the use of rubric: in a few cases when the rubric was used but not declared to students, they started to ask about the evaluation criteria.
- The studio team started to feel the need to have a rubric for the projects conducted in the Q2, Q3 and Q4 periods. Having become familiar with the logic, requirements and structure of the rubric, they could generate specialized ones for each project very quickly.

The criterion *Problem Solution* refers to how well the problem in the given brief is addressed visually. In the Q1 period, the brief usually asks for a composition that emphasises a specific design principle. For example, when we ask students to make a composition that emphasizes contrast, this criterion evaluates how well contrast is presented in the composition, from non-existent to a creative visual communication of contras.

The criterion Organisation refers to the general organisation of the composition, from randomly scattered elements to a dynamic order.

Figure 5. Proposed Rubric (visual generated by the authors).

	0-8 pts.	9-11 pts.	12-14 pts.	15-17 pts.	18-20 pts.
PROBLEM SOLUTION	Shows no solution	Adopts a conceptual or schematic approach	Develops a partial visual solution	Develops a full visual solution	Develops a creative, novel solution
ORGANIZATION	Uses elements unconsciously and randomly	Establishes order among some of the elements	Established symmetrical order	Establishes a non-symmetrical order	Forms an aesthetic and dynamic order
CHOICE AND USE OF ELEMENTS	Chooses inconsistent and inharmonious elements	Uses a single type of element (in terms of form, color or size)	Uses a variety of elements in harmony	Uses a variety of elements in multiple harmonious groups	Relates groups of elements on multiple layers
CONSTRUCTING KNOWLEDGE	Shows no reference to any compositional fundamentals	Shows some references and certain visual relationships	Configures the composition by using learned principles	Applies learned principles for a compositional purpose	Configures multiple principles and use them creatively
DEFINITION AND USE OF AREA	Doesn't define the area	Ill defines the area	Defines the area without contributing to the composition	Defines the area in service of the composition but ignores negative spaces	Defines the area and forms meaningful negative spaces that add value to the composition

The criterion Choice and Use of Elements refers to how well the visual elements are chosen for the given topic and if a variety of elements are used rather than only one type or one size of elements, which undeniably makes the organisation more challenging. So, this criterion helps assess whether the elements used are just randomly different in certain attributes or are purposefully varied in accordance with the organisational idea.

The criterion Constructing Knowledge refers to the application of compositional fundamentals - such as aligning elements to create a visual relation, arranging distances or applying systematic change - onto the current work. This criterion helps assess how well students build upon their learning, from showing no reference to any fundamentals of composition to creatively using multiple principles together.

The criterion Use and Definition of area refers to compositional space and how it is used in terms of placing elements on the given area. The hierarchy levels start from being unaware of the spatial constraints to fully making use of them in service of the composition.

CONCLUSION

In this study we explained our attempt to systematize assessment in the foundation design studio. Assessment has rarely been approached as a topic of academic inquiry for the first year of design education (Demirkan and Afacan, 2012) and we hope that our initial attempt contributes to a change in our field. The evaluation of our rubric by both experts and studio tutors confirms that this systematic assessment method is easier, quicker and does more justice to student works as it allows for a more in depth and varied analysis of individual student works, compared to traditional approaches that rely on comparison and discussion.

Beyond these immediate benefits, the most important aspect of a rubric is the consistency it provides for assessment as verified by our inter-rater reliability analysis. It would act as a guide for outside or novice educators in assessment, such as invited jury members or teaching assistants, and ensure more consistent and valid results. Indeed, less experienced tutors and teaching assistants expressed feeling more comfortable while evaluating with the provided rubric. However, it still leaves room for expert evaluation, as an example “using a variety of elements” can be a good thing if composed well with clear relations among the elements, yet it can also be devaluing for the composition if these elements are used randomly, which still needs expert evaluation to differentiate. Moreover, even if the rubric is not fully developed or does not end up being used, just the exercise of developing it brings the instructor team closer in understanding and approach of assessment. Based on our own experience, we can confidently state that while trying to define the criteria and their levels of complexity, teachers re-examine their own evaluation process, understand when or why their perception or judgement might sometimes be altered, and remember to re-focus on the intended course learning outcomes.

In this study we focused on the Q1 period of our foundation design studio course aimed at teaching basic design elements and principles. The works produced in this period are 2D artefacts made out of paper and take a few hours to make. To benefit the whole foundation year with such systematic assessment, criteria and competence levels for the assessment of other assignments dealing with different materials, 3D forms, longer processes

with multiple iterations and other concerns needs to be developed. For example, problem solution and organisation are valid for all assignments of the foundation studio course but definition and use of area is specific to 2D works, which may be replaced with exploration of 3D space in Q3. Furthermore, we have yet to explore how to assess other intended learning outcomes of our course such as being self-critical in the editing of one's work and using a vocabulary of terms specific to the design activity. For projects beyond the 2D works of the Q1 period studied here, some of the assessment criteria can be changed and the rubric can still be used. For example, Problem Solution might change to Approach to Design Brief, or Use and Definition of Area might change to Use and Definition of Constraints.

The Revised Rubric was statistically validated through quantitative measures; however we did not extend our study to explore the reasons behind raters decisions. The current study can be followed up with qualitative methodologies to explore more in depth the grading process and reasoning behind assessment of 2D compositions. It will also be useful to delve deeper into outliers, where one expert assessed very differently on one criterion than others, to explain the underlying reasons and to open more room for expert evaluation if necessary. Similarly, a closer inspection of cases where experts evaluated differently on different criteria but arrived at a more or the less similar end grade worth exploring how and why they differentiate on certain criteria.

Scholars have already documented the positive effects of rubrics on students' learning, which is especially important for the foundation studio where students often voice their confusion on what is expected, how they should achieve it, and how they are assessed. During the three terms when the rubric was used to assess student works as part of the field testing, there were less concerns and questions from students regarding their results or our assessment approach than ever before. It should be noted that, since we focused on the first period of the first term, our approach to rubric development was "top-down" rather than a "bottom-up" (Nitko and Brookhart, 2011) where students participate in the development and use of rubrics. Since studies suggest that the bottom-up approach results in improvements in academic performance (Reitmeier et al., 2004; Petkov and Petkova, 2006), involving students in the structuring, wording and application of the rubric can be explored in future studies, which will be appropriate for the constructivist learning environment of the studio. Another point that the constructivist pedagogy requires is the assessment of the student's process and development, which were left out in this initial study but should be a major topic of exploration in the future.

Finally, during this process we realised that the intrinsic connection between the assessment tasks, criteria, intended learning outcome, and teaching and learning activities, i.e., what Biggs and Tang (2011) call the constructivist alignment in education, calls for a potential revision of both the learning outcomes and the activities of our course. We have studied examples of constructive alignment applied to CAD courses (Gelmez, and Arkan, 2022) or structures courses (Acican and Luyten, 2025) and we call for a revision of basic design in a similar way.

APPENDIX A: The First Draft of the Rubric

	poor	average	good	very good	exceptional
Solution to problem	None	The problem is understood but not solved	A basic, simple solution	A more complex solution	Creative solution
Proof of visual consciousness or order	None	Some connection between elements	A legible order but boring (symmetric)	More interesting order with some mistakes (asymmetric)	Very good order
Execution and presentation (format, craft, colour)	Poor	An initial awareness	Sensitivity observed	Refined with some small mistakes	Purposeful use of colour, perfect craftsmanship
Design principles	None	A sense of knowledge but not used properly	Somewhat used	Applied very clearly	Used in unexpected manner

APPENDIX B: The Tentative Rubric

	0-8 pts.	9-11 pts.	12-14 pts.	15-17 pts.	18-20 pts.
PROBLEM SOLVING	no solution	diagrammatic of symbolic approach	visual solution attempted	Visual solution achieved	Creative solution
OVERALL STRUCTURE	elements placed randomly or combined into a single unit	some order OR symmetrical arrangement OR pattern	Symmetrical order with a tilt or shift of axis OR attempted asymmetrical order	A non-symmetrical order is achieved	Aesthetic order
CHOICE AND USE OF ELEMENTS	inconsistent and inharmonious choice and placement of elements	Only one type of element (with or without size/colour difference) OR variety of elements with little coherence	Variety of elements used coherently	Variety of elements put in relationship through groupings	Groups of elements put in relationship in multiple layers
ACQUIRED KNOWLEDGE /DESIGNERLY ATTENTION	No prior knowledge is conveyed	Demonstrates some conscious decisions (attempt at aligning OR distancing OR purpose, such as direction)	Designerly sensitivity in arrangement of elements OR expression of purpose	Designerly sensitivity in arrangement of elements AND expression of purpose	Principles and acquired knowledge used together in an innovative manner
DEFINITION AND USE OF AREA	Area is not defined, OR ill defined	Area defined with primitive methods such as creating a frame, filling the corners, etc.	Organization of the elements is forced to fit the area OR area definition is not fully working	Area defined in a natural (unforced) manner through the organization of the elements	Area and its negative spaces are used in a way that adds value and meaning to the overall composition

APPENDIX C: The details of the Revised Rubric used in the field testing, as seen in Google Classroom

Problem Solving				
0-8 pts	*9-11 pts*	*12-14 pts*	*15-18 pts*	*18-20 pts*
Shows no solution	Adopts a conceptual or schematic approach	Can partially solve the problem visually	Can fully solve the problem visually	Can solve the problem visually in a creative, novel way

Organization				
0-8 pts	*9-11 pts*	*12-14 pts*	*15-18 pts*	*18-20 pts*
Uses elements unconsciously and randomly	Can establishes order among some of the elements	Can establish symmetrical order	Can establish a non symmetrical order	Can form an aesthetic and dynamic order

Choice and use of elements				
0-8 pts	*9-11 pts*	*12-14 pts*	*15-18 pts*	*18-20 pts*
Chooses inconsistent and inharmonious elements	Uses a single type of element (form, color, size)	Can uses variety of elements in harmony	Can use variety of elements in multiple harmonious groups	Can relate groups of elements on multiple layers

Use of acquired knowledge and principles				
0-8 pts	*9-11 pts*	*12-14 pts*	*15-18 pts*	*18-20 pts*
No prior knowledge is conveyed	Shows traces of prior knowledge	Can constructs work using learned principles	Can apply learned principles for a compositional purpose	Can configure multiple principles and use them creatively

Definition and use of area				
0-8 pts	*9-11 pts*	*12-14 pts*	*15-18 pts*	*18-20 pts*
Area not defined, or ill defined	Defines area with improper or primitive methods (creating frame, filling corners, etc.)	Defines area by forcing element placement and ignoring negative spaces	Can define and use the area naturally through the organization of elements	Can define the area well and form meaningful negative spaces that add value to the composition

REFERENCES

ACAR, A., SOYSAL ACAR, A. Ş., ÜNVER, E. (2019) Mimarlık Bölümü Birinci Sınıf Öğrencilerinin Görsel-Mekânsal Becerileri Üzerine Bir Araştırma, *METU Journal of the Faculty of Architecture* 36(2) 73-92.

ACICAN, O., LUYEN, L. (2025) Experiential Learning of Structures through Dynamic Comparisons: A Comparative Study, *International Journal of Technology and Design Education*, 1-26.

AFLATOONY, L., WAKKARY, R., NEUSTAEDTER, C. (2018) Becoming a Design Thinker: Assessing the Learning Process of Students in a Secondary Level Design Thinking Course, *International Journal of Art & Design Education* 37(3) 438-453.

ANDRADE, H., DU. Y. (2005) Student Perspectives on Rubric-referenced Assessment, Practical Assessment, *Research & Evaluation* 10(5) 1–11.

ARI, A. (2013) Bilişsel Alan Sınıflamasında Yenilenmiş Bloom, SOLO, Fink, Dettmer Taksonomileri ve Uluslararası Alanda Tanınma Durumları, *Uşak Üniversitesi Sosyal Bilimler Dergisi* 6 (2) 259-290.

ARKUN KOCADERE, S., ÖZGEN, D. (2012) Assessment of Basic Design Course in Terms of Constructivist Learning Theory, *Procedia - Social and Behavioral Sciences* 51 115-119.

BARTHOLOMEW, S. R., RUESCH, E. Y., HARTELL, E., STRIMEL, G. J. (2020) Identifying Design Values Across Countries Through Adaptive Comparative Judgment, *International Journal of Technology and Design Education* 30(2) 321-347.

BIGGS, J. (1999) What the Student Does: Teaching for Enhanced Learning, *Higher Education Research and Development* 18 57-75.

BIGGS, J. (2013) *Constructive Alignment in University Teaching*, HERDSA Review of Higher Education 1.

BIGGS, J., TANG, C. (2011) *Teaching for Quality Learning at University: What the Student Does*, Open Berkshire, University Press.

BINGHAM, R. (2002) Planning a Unit of Study. [http://ctet.royalroads.ca/writing-effective-assessment-criteria] Access Date (08.07.2021)

BOHEMIA, E., HARMAN, K., McDOWELL, L. (2009) Intersections: The utility of an 'Assessment for Learning' Discourse for Design Educators, *Art, Design and Communication in Higher Education* 8(2) 123-134.

BOLING, E., SCHWIER, R.A. (2016) Curators Notes,, *Studio Teaching in Higher Education: Selected Design Cases*, der. E. Boling, R.A. Schwier, C.M. Gray, K.M. Smith, K. Campbell, Routledge.

BROOKHART, S.M. (2013) *How to Create and Use Rubrics for Formative Assessment and Grading*, Alexandria, VA: ASCD.

CAMPBELL, A. (2005) Application of ICT and Rubrics to the Assessment Process here Professional Judgment is Involved: The Features of an E-marking Tool, *Assessment & Evaluation in Higher Education* 30(5) 529–537.

CROWTHER, Phillip. (2013). Understanding the Signature Pedagogy of the Design Studio and the Opportunities for its Technological Enhancement. *Journal of Learning Design*, 6 (3), 18-28.

ÇIKIŞ, Ş., ÇIL, E. (2009). Problematization of Assessment in the Architectural Design Education: First Year as a Case Study. *Procedia Social and Behavioral Sciences*, 1, 2103-2110.

ÇIL, E., PAKDIL, O. (2017). Design Instructor's Perspective on The Role of Computers In Architectural Education: A Case Study, *METU Journal of the Faculty of Architecture*, 24(2), 123-136

DAVIES, A. (2003). ADC- LTSN Learning and Teaching Fund Project: Effective Assessment in Art and Design. Retrieved July 8, 2021, from http://www.brighton.ac.uk/

DAVIES, A. (2012). Learning Outcomes and Assessment Criteria in Art and Design. What's the Recurring Problem?. Networks 18. Retrieved July 8, 2021, from http://arts.brighton.ac.uk/

DE LA HARPE, B., PETERSON, J. F., FRANKHAM, N., ZEHNER, R., NEALE, D., MUSGRAVE, E., MCDERMOTT, R. (2009). Assessment Focus in Studio: What is Most Prominent in Architecture, Art and Design?. *International Journal of Art and Design Education*, 28 (1), 37-51.

DEMIRKAN, H., AFACAN, Y. (2012). Assessing creativity in design education: Analysis of creativity factors in the first-year design studio. *Design Studies*, 33(3), 262-278.

GELMEZ, K., ARKAN, S. (2022). Aligning a CAD course constructively: telling-to-peer and writing-to-peer activities for efficient use of CAD in design curricula. *International Journal of Technology and Design Education*, *32*(3), 1813-1835.

GASAYMEH, A. M. (2011). The Implications of Constructivism for Rubric Design and Use [Conference presentation]. Higher Education International Conference (HEIC 2011). http://heic.info/assets/templates/heic2011/papers/05-Al-Mothana_Gasaymeh.pdf Access Date (5.7.2021).

GILOI, S., DU TOIT, P. (2013). Current Approaches to the Assessment of Graphic Design in a Higher Education Context. *The International Journal of Art and Design Education*, 32 (2), 256-268.

GISEV, N., BELL, J.S., CHEN, T.F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy*, 9(3), 330-338.

GREGORI GIRALT, E., JOSÉ LUIS MENÉNDEZ-VARELA, J.L. (2012). Performance assessment and rubrics in art education: A study of student perceptions. Observar, 12, 39-52. https://dialnet.unirioja.es/servlet/articulo?codigo=7448997 Access Date (5.12.2021).

GROBLER, A. (2006). The Role of Assessment in the Design Process [Conference presentation]. DEFSA Conference. http://www.defsa.org.za/node/276 Access Date (8.7.2021).

GÜR, B. F., YÜNCÜ, O. (2010). An integrated pedagogy for 1/1 learning. *METU Journal of the Faculty of Architecture. (27:2) 83-94.*

GVILI, I.E., WEISSBURG, M.J., YEN, J., HELMS, M.E., TOVEY, C.A. (2016). Development of scoring rubric for evaluating integrated understanding in an undergraduate biologically-inspired design course. *International Journal of Engineering Education*, 32, 123-135.

HASDOGAN, G. (2012). Characterising Turkish Design Through Good Design Criteria: The Case Of 'Design Turkey' Industrial Design Awards. *METU Journal of the Faculty of Architecture*, 29(1), 171-191.

IORDANOVA, I. (2007). Teaching Digital Design Exploration: Form Follows… *International Journal of Architectural Computing*, 4(5) 686-702.

JONSSON, A., SVINGBY,G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144.

KARAHANOĞLU, A. (2022). Act like a user, work like a teacher: evaluation of experience design works through peer-testing. *International Journal of Technology and Design Education*, *32*(5), 2827-2852.

KARAHANOĞLU, A., OUDE ALINK, C., BAKIRLIOĞLU, Y. (2019). Quantifying Design for User Experience Assignments: Using Rubrics as Assessment Tools [Conference presentation]. Proceedings of DRS Learn X Design 2019: Insider Knowledge - Fifth International Conference for Design Education Researchers.

KETIZMEN, G., GÜÇYETER, B. (2024). Effect of first-semester basic design studio on architecture students' creative skills: a pre-and post-test assessment. *International Journal of Technology and Design Education*, *34*(4), 1603-1655.

KETIZMEN, G., KELEŞ, H. (2025). Development and validation of a design process assessment scale for architectural design studios. *International Journal of Technology and Design Education*, 1-25.

LEVENT KASAP, T., KAPTAN, B. B. (2022). Design studio final product evaluation rubric in interior architecture education: Eskişehir Technical University case. *IDA: International Design and Art Journal*, 4(2), 185-199.

LEUNG, C. F. (2000). Assessment for Learning: Using SOLO Taxonomy to Measure Design Performance of Design and Technology Students. *International Journal of Technology and Design Education*, 10, 149-161.

NUHOĞLU, K., AKKOYUNLU, P., AKKOYUNLU, B. (2017). Fostering and assessing infographic design for learning: the development of infographic design criteria. *Journal of Visual Literacy*, 36(1), 20-40.

ONWUAGBOKE, B. B. C., SINGH, T. K. R. (2016). Reliability and validity of graphic design assessment rubrics. *International Journal of Technical Research and Applications*, 4(2), 119-124.

ORR, S., BLOXHAM, S. (2012). Making Judgements about Students Making Work: Lecturers' Assessment Practices in Art and Design. *Arts and Humanities in Higher Education*, 12(2–3), 234–253.

ÖZGEN KOÇYILDIRIM, D., TONUK, D. (2020). Değişimi Ölçmek: Temel Tasarım Projeleri için Değerlendirme Ölçütleri. *UTAK2020 4. Ulusal Tasarım Araştırmaları Konferansı*, Ankara, Turkey, 8 - 10 September 2020, pp.275-285.

PANADERO, E., JONSSON, A. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles Policy and Practice*, 21(2),133-148.

PANADERO, E., JONSSON, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*.

PARKES, K. A. (2010). Performance Assessment: Lessons from Performers. *International Journal of Teaching and Learning in Higher Education*, 22(1), 98–106.

PETKOV, D., O. PETKOVA. 2006. Development of Scoring Rubrics for IS Projects as an Assessment Tool. *Issues in Informing Science and Information Technology* 3: 499–510.

POP-ILIEV, R., PLATANITIS, G. (2008). A rubrics-based methodological approach for evaluating the design competency of engineering students. Proceedings of the TMCE 2008.

RESNICK, M. (2017). *Lifelong Kindergarten: Cultivating Creativity through Projects, Passions, Peers, and Play*. MIT Press.

ROHSE, S., ANDERSON, T. (2006). Design Patterns for Complex Learning. *Journal of Learning Design*, 1(3), 82-91.

ROY, B., BHATTACHARYYA, T. & BANERJEE, P. (2022). Active Learning in Higher Education by SOLO Taxonomy. Proceedings of the International Conference on Best Innovative Teaching Strategies ICON-BITS 202. doi:dx.doi.org/10.2139/ssrn.4022499

RUST, C., PRICE, M., O'DONOVAN, B. (2003). Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes. *Assessment and Evaluation in Higher Education*, 28(2), 147-164.

SEERY, N., CANTY, D., PHELAN, P. (2011). The Validity and Value of Peer Assessment Using Adaptive Comparative Judgement in Design Driven Practical Education. *International Journal of Technology and Design Education*, 22, 205–226.

SCHELLING, J., LEURS, B., & BEST, S. (2012). From gut feeling to a structured, summative assessment of Design competencies. In DS 74: *Proceedings of the 14th International Conference on Engineering & Product Design Education (E&PDE12) Design Education for Future Wellbeing, Antwerp, Belgium*, 06-07.9. 2012 (pp. 322-327).

SCHULMAN, L. S. (2005). Signature Pedagogies in the Professions. *Dedalus*, 134, 52-59.

SCHON, D.A. (1983). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York.

SMITH, K. M. (2013). Assessment as a Barrier in Developing Design Expertise: Interior Design Student Perceptions of Meanings and Sources of Grades. *International Journal of Art and Design Education*, 32(2), 203–214.

SOLECKI, I., VITOR PORTO, J., DA CRUZ ALVES, N., VON WANGENHEIM, C.G., HAUCK, J.C.R., FERRETI BORGATTO, A. (2020). Automated Assessment of the Visual Design of Android Apps Developed with App Inventor. Proceedings of the 51st ACM Technical Symposium on Computer Science Education, 51–57.

STEMLER, S. (2004). A Comparison of Consensus, Consistency, and Measurement Approaches to Estimating Interrater Reliability. *Practical Assessment, Research, and Evaluation*, 9, 1-19.

TABER, K.S. (2017). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48:1273–1296.

TASTAN, H., ER, İ.E. (2025). A systematic approach to evaluating architectural portfolios*:* architectural portfolio evaluation rubric (APER). *International Journal of Technol Design Education* 35, 2093–2128.

SUDHINDRA, S. T., BLESSING, L. T. (2021). A framework for design competency assessment. *Proceedings of the Design Society*, 1, 91-100.

UZUNOĞLU, K. & UZUNOĞLU, S. (2011). Project Evaluation Process with Classified Objective Criteria in Architectural Education. *Procedia - Social and Behavioral Sciences*, 28, 1004-1010.

WEBSTER, H. (2007). The Assessment of Design Project Work (Summative Assessment). EBE Briefing Guide Series 9. from http://www.heacademy.ac.uk/ Access Date (8.1.2021).

WOODHALL, T.F.C. (2008). Redesigning Assessment: The Design and Implementation of a Rubric-based Assessment System to Improve Engineering Design Education [Unpublished doctoral dissertation]. Queen's University.

YUAN, H., WU, Y., TAO, H., YIN, J., FANG, Y., ZHANG, J., ZHANG, Y. (2025). Construction of a sustainable design competency assessment system for fashion designers in China. *International Journal of Technology and Design Education*, *35*(1), 305-332.

Alındı: 12.05.2025; **Son Metin:** 20.04.2026

Anahtar Sözcükler: Tasarım eğitimi; temel tasarım; rübrik; tasarım stüdyosu için rübrik; stüdyo pedagojisi.

TEMEL TASARIM STÜDYOSUNDA ÖĞRENCİ ÇALIŞMALARININ DEĞERLENDİRİLMESİ İÇİN RUBRİK GELİŞTİRİLMESİ

Bu çalışma temel tasarım dersinde öğrenci projelerinin değerlendirilmesini iyileştirmek adına bir rübrik geliştirilmesini önermektedir. Bu tasarım eğitiminde değerlendirme konusunda ihtiyaç duyulan sistematizasyon ve teorizasyonu sağlayabilmek açısından önemlidir. Bunu yapabilmek için öncelikle tasarım eğitimi ve bu eğitime uygun değerlendirme stratejileri tartışılacaktır. Daha sonra bu çalışmada uygulanan, portfolio denemesi, notlama matriksleri, deneyimli eğitimcilerle görüşme ve bunların istatiksel tutarlılık testi ve 3 dönemlik saha denemesi aşamalarından oluşan kapsamlı metod anlatılacaktır. Bu çalışmaların sonucunda geliştirilen rübriğin tasarım eğitimi için işlerliğinden bahsetmek mümkündür ve ayrıca rübrik kullanımının değerlendirmeyi daha tutarlı, verimli, derin ve adil yaptığı sonucuna varılmıştır. Böylelikle bu çalışma hem rübrik geliştirmek isteyen eğitimcilere yol gösterici olurken hem de geliştirilen rübriği kullanmak veya uyarlamak isteyen eğitimcilere sunar.

DEVELOPING A RUBRIC TO ASSESS STUDENT WORKS IN FOUNDATION DESIGN COURSE

This study aims to improve assessment in foundation design studio courses by bringing much-needed systematization and theoretical grounding so that learning goals are better communicated among the tutors and with the students. For this purpose, first, the specific pedagogy of foundation design studio courses is detailed to better explain the need for specifically tailored assessment criteria. Accordingly, a review of literature is presented on the ways to define and measure assessment criteria with an assessment rubric based on Biggs' SOLO taxonomy. A method consisting of 5 consecutive stages is used to develop this rubric for our specific context: 1- definition of assessment criteria, 2- ranking matrices and portfolio testing, 3- expert interviews, 4- inter-rater reliability analysis, 5- field testing for three

academic terms. Based on the outcomes of these stages, we argue in favor of using rubrics for assessment in foundation design studios, since this approach makes grading more consistent, time-efficient, in-depth and fair. Moreover, the rubric development process in itself was found to be very beneficial for the tutor team as it unified their approach to education and assessment. This study both provides a guideline for the educators who want to develop their own rubric, as well as presenting the rubric we developed to improve assessment in foundation design studios.

DALSU ÖZGEN KOÇYILDIRIM; BFA, MFA, DFA
Received her bachelor's degree on graphic design from Bilkent University, master's degree on design and technology from Parsons School of Design, DFA on Graphic Design from Mimar Sinan Fine Arts University. Her research interests include basic design education, design communication and visual narratives. dozgen@metu.edu.tr

DAMLA TONUK; BSc., MSc., PhD.
Obtained bachelor's and master's degrees from the Department of Industrial Design at METU, and PhD from the Department of Sociology, Lancaster University, UK. Her work centers around applying material culture and science and technology studies perspectives to understanding human-object-material relationships. dtonuk@metu.edu.tr