Chapter 10
Developing a Rubric to Evaluate the Dissertations Conducted in the Fields of Educational and Social Sciences

Ömer Açıkgöz
Social Sciences University of Ankara, Turkey

Aydın Aslan
Selcuk University, Turkey

Korkut Koçak
Republic of Türkiye Ministry of National Education, Turkey

Aslı Günay
Social Sciences University of Ankara, Turkey

Nevzat Yavuz
The Scientific and Technological Research Council of Türkiye, Turkey

ABSTRACT

This study aims to develop a rubric to evaluate the dissertations implemented in the fields of educational and social sciences. In the development of this rubric, the acquisition requirements concerning knowledge, skills, and competence at the doctorate level in the European and Turkish qualifications frameworks, the legal framework of Turkish higher education, and the perceptions of 12 experts in the fields of educational and social sciences concerning the common competences of the dissertations were considered. The rubric can contribute to the evaluation of dissertations completed in the field of educational and social sciences concerning these dimensions and provide PhD students, researchers, and academics with a guide to evaluate their academic studies based on an empirical instrument.

DOI: 10.4018/978-1-6684-6086-3.ch010

INTRODUCTION

A Doctor of Philosophy (Ph.D.) dissertation is the final and highest educational outcome in higher education. Hall (1991) indicated that the academic significance of a discipline depends on the quality of doctorate studies, in which Ph.D. dissertations are seen as important tools for a discipline to improve academically. Briefly, Ph.D. dissertation demonstrates its author's technical, analytical, and writing skills (Lovitts, 2005). The doctorate level is the highest qualification level in the European Qualifications Framework (EQF) and the Turkish Qualifications Framework (TQF), which refers to level 8 and usually requires four years of study, mostly as a period of research. Ph.D. holders are expected to confirm their advanced knowledge, skills, and competencies in their dissertations according to their qualification level description (EQF, 2015; TQF, 2016). Besides, national authorities can take additional measurements to improve the quality of the doctorate programs and Ph.D. dissertations. For example, the Turkish Graduate Education Regulation (TGER) in 2016 emphasizes that the Ph.D. holders in Türkiye are required to meet one of the conditions, namely bringing innovation to science, developing a new scientific method or implementing a known method to another field (TGER, 2016). In this regard, they are expected to contribute to the literature or the application through their dissertations. They are important scientific studies with findings, implications, and suggestions attracting various researchers, policymakers, and academics.

Several questions appear in the evaluation of a dissertation: what determines a basic dissertation with minimum requirements? Who decides what characteristics it should have? These are the critical questions, which should be considered in the evaluation of Ph.D. dissertations. In particular, although countries have different higher education systems and regulations, most universities have issued some general or specific guidelines on quality and standards for Ph.D. dissertations evaluations regardless of the locations. For instance, properties like originality, sound methods, significant contribution to knowledge, and publishable results are common worldwide criteria (Kyvik & Thune, 2015).

While the TGER establishes a standard framework for evaluating Ph.D. dissertations in Türkiye, each university in Türkiye has its own regulations for Ph.D. dissertations processes, regardless of the academic field. The dissertation jury consists of five members: three faculty members, including the student's dissertation supervisor, and at least two representatives from other higher education institutions. Moreover, evaluation processes might differ according to the students' and supervisors' characteristics, experiences, and relations. Hence, it is not possible to say that they have particular criteria to evaluate dissertations in Türkiye. For that reason, this study tries to fill this gap in the literature by developing a rubric through which jury members, academics, researchers and doctoral students can evaluate the dissertations conducted in the fields of educational and social sciences. Furthermore, the dimensions in the rubric can provide them with a profound and comprehensive idea about the quality of the Ph.D. dissertations.

Developing a rubric to evaluate Ph.D. dissertations conducted in the fields of educational and social sciences based on the doctorate qualifications, based on knowledge, skills, and competencies descriptors can contribute to the literature. Regarding this, qualification descriptors of the EQF and TQF were used, providing a comprehensive overview of the quality of doctorate programs and Ph.D. dissertations in European countries and Türkiye for the possible acquiring learning outcomes during their studies. In this regard, universities in both Türkiye and European countries might use the newly developed rubric by this study as a reference source to evaluate the quality of Ph.D. dissertations.

BACKGROUND

The rubric, which is often used in dissertation evaluations, is one method of evaluating dissertations. It is a scoring tool used in qualitative rating to evaluate original or complex student works. It consists of criteria for rating significant dimensions of performance and standards to what extent those criteria are attained (Jonsson & Svingby, 2007). It promotes peer assessment, self-assessment, and academics assessment (Andrade, 2005). In this regard, academics, educators, and researchers can use it with higher education students. There are three fundamental features in a rubric: evaluation criteria, quality dimensions, and a scoring strategy (Popham, 1997). It can decrease rater biases with its construct consisting of criteria and performance levels and provide students with more realistic and detailed feedback concerning their performance (Parlak & Doğan, 2014).

The rubric can be holistic or analytic, depending on its purpose of usage. Scoring in a holistic rubric is implemented based on students' overall impressions. It gives evaluators a quick scoring and overall impression of students' performance on a task despite not involving detailed feedback. On the other hand, scoring in an analytic rubric is conducted according to several dimensions. Therefore, evaluators can get more detailed feedback and consistent scoring across students and graders though it is time-consuming to score (Zimmaro, 2004). Another important distinction about the rubric is that it can be general and task-specific. A general rubric is used when all performances and tasks are evaluated with the same rubric. If a particular performance or task is aimed to be evaluated, a task-specific rubric is employed (Özen, 2019). The purpose of usage determines whether a rubric will be holistic, analytic, general or task-specific.

Rubrics are employed in many disciplines such as the liberal arts, management, and teacher education (Reddy & Andrade, 2010). For example, a rubric was developed to grade undergraduate students' APA-style introductions (Stellmack, Konheim-Kalkstein, Manor, Massey & Schmitz, 2009). Rubrics are consulted to evaluate doctoral dissertations (Agu, Omenyi & Odimegwu, 2015; Fitt, Wlaker, Leary, 2009; Lovitts, 2005). One reason to use rubrics in the evaluation of doctoral dissertations is that they make the evaluation criteria more explicit (Johnsson & Svingby, 2007). Therefore, several universities (Texas A & M Commerce University, 2016) utilize them to evaluate theses and dissertations.

In this context, the main objective of this study is to develop a rubric to evaluate the Ph.D. dissertations implemented in the fields of educational and social sciences. Hence, standard criteria might be defined for the evaluation processes of Ph.D. dissertations to improve their quality. As the qualifications in the EQF, TQF, and TGER consist of several dimensions, an analytic rubric was embraced to elicit more detailed feedback for the dissertations completed in the field of educational and social sciences with regard to theoretical framework, method, and contributions of research dimensions. The knowledge, skills, and competencies of researchers can be measured in their dissertations with this rubric. For this purpose, 20 dissertations were randomly selected from the National Theses Center of the Council of Higher Education database in Türkiye, 10 in the field of educational sciences and 10 in the field of social sciences. Besides, 12 experts in different educational and social sciences departments in Turkish universities (two experts in measurement and evaluation in educational sciences, one professor in educational administration, two professors in educational sciences, one associate professor in psychology, one associate professor in sociology, one associate professor in mathematics, one assistant professor in economy, one assistant professor in international relations, one professor in economy, and one expert in public administration) were asked to define the common criteria for Ph.D. dissertations to lessen the quality differences among them in the fields of educational and social sciences.

METHODOLOGY

In developing the current rubric, the dimensions with regard to the EQF, TQF, and TGER were considered. Besides, the studies conducted by Karslı, Karabey, Çağıltay, and Göktaş (2018) and Lovitts (2005) were taken into account qualification dimensions to determine the items in the rubric. The perceptions of 12 experts in their fields concerning their academic standards were also taken into account to form the items in the rubric. The rubric consists of three dimensions "Theoretical Framework", "Method" and "Contributions of Research (AppendixA). The first one is related to the clarity of the scientific importance, research problem, and the theory/theories on which the dissertation is based. The second one is closely related to scientific knowledge and appropriate methodologies, including the notions of validity, reliability, correctness, truthfulness, and consistency. The last one indicates originality, scientific contributions to knowledge, publishable results, and practical and socio-economic assistance.

After all, these dimensions were formed based on the EQF and TQF, TGER, the literature review and the perceptions of 12 experts in the field of educational, and social sciences. Hence, there are 11 items in the rubric, which describe these dimensions. The items are rated in accordance with "no", "partial" and "yes". A three-category grading was chosen in this rubric because of two reasons. The first reason is that a gradual and hierarchic three-category grading in the rubric makes the criteria of the rubric more understandable and explicit. The second one is that it was decided that using a three-category grading would be more appropriate based on the perceptions of 22 experts (seven experts in the field of measurement and evaluation, one associate professor, and three assistant professors in the field of educational administration, two assistant professors in the field of educational sciences, two associate professors and one assistant professor in curriculum and instruction, one associate professor in sociology, one assistant professor in psychology, one professor and one assistant professor in the economy, one expert in the field of public administration and one assistant professor in the international relations) concerning the items in the rubric.

In the development of the rubric, the perceptions of 22 experts considered. It was paid attention to select the experts with the theoretical and practical knowledge to be able to supply more comprehensive and profound feedback for the rubric. 15 in 22 experts have been experienced to be a jury member in master's thesis and doctoral dissertation defenses.

The researchers of this study in collaboration with an independent researcher carefully examined the dimensions in the EQF, TQF, and TGER, in addition to the literature review and 12 experts' perceptions concerning the common criteria for the high-quality dissertations to determine which items should be included under each dimension in the rubric. In this context, 15 tentative items were prepared and sent to 22 experts. These experts were sent an evaluation form to indicate their perceptions of the items and were asked to score the tentative 15 items in the rubric as "Acceptable", "Revised" and "Not acceptable".

In the context of the content validity of the tentative data collection instrument, Content Validity Ratio (CVR) and Content Validity Index (CVI) based on expert views (Lawshe 1975, 568) were used to determine which items should be omitted in the rubric. In the determination of experts, who have adequate knowledge and enough time to evaluate a data collection instrument, at least 5 experts and at most 40 experts can take part in an evaluation of the instrument (Yurdugül 2005, 2). In the current rubric development, the perceptions of 22 experts were derived in the calculation of CVR for each item and CVI for the rubric and its dimensions. CVR is calculated in the following way (Lawshe 1975, 567):

$$CVR = N_E - (N/2) / N/2$$

N_E shows the number of experts indicating an item "Acceptable" and N the number of experts in this formula. The items with 0 (zero) or lower than this value are directly omitted from the tentative instrument and Content Validity Criterion (CVC) is calculated for each remaining item at the level of significance (α=.05). CVC (CVRcritical) is a minimum CVR value required to remove the possibility to find each item appropriate in the instrument by chance and make correct decisions for items.

FINDINGS

According to CVRcritical the table prepared by Ayre and Scally (2014) in consideration of Lawshe's study, the minimum CVR value should be at least 0.455 for one item 22 experts examine at the level α=0.05 (type 1 error). According to Table 1, 11 in 15 items met the accepted conditions (CVR ≥ CVR critical (N=22. α=0.05)=.455).

Table 1. CVR values for the tentative items

Item	CVR		Item	CVR
Item1	1,000		Item9	0.158*
Item2	0,895		Item10	0.579
Item3	0,789		Item11	0.684
Item4	0,684		Item12	-0.158*
Item5	0,895		Item13	0,789
Item6	0,789		Item14	0.684
Item7	0,053*		Item15	0.368*
Item8	0,895			
*CVR<$CVR_{Critical}$(N=22, α=.05)				

Content Validity Index (CVI) is calculated for all tests and if available its sub-dimensions after CVR treatment to include items in the final test. CVI critical value is the same as CVR, which is valid for items (Yurdugül, 2005). So, the minimum CVI value for the rubric and its sub-dimensions is .455 for 22 experts at the level of α=0.05 (type 1 error).

Table 2. CVI values for all rubric and its sub-dimensions

Rubric and its sub-dimensions		CVI
The Whole/entire rubric		.789*
Theoretical framework		.895*
Method		.789*
Contribution of research		.726*
*CVI≥$CVI_{critical}$(N=22, α=.05)		

According to Table 2, CVI values for the rubric and its sub-dimensions are higher than CVI critical value (CVI ≥ CVcritical (N=22, α=0.05)=.455). Based on CVR and CVI values, the content validity of the rubric is acceptable.

Besides, the construct validity of the rubric was ensured through the perceptions of these experts to evaluate how well the rubric measures dissertations with regard to theoretical framework, method and contribution of research dimensions. Its reliability was checked through the rater reliability because this type of reliability is the most frequently considered type of reliability in rubric development (Moskal & Leydens, 2000). So, 20 dissertations were randomly selected from the National Theses Center of the Council of Higher Education database, 10 in the field of educational sciences and 10 in the field of social sciences. Six experts in their disciplines, three in educational sciences and three in social sciences assessed them. All the experts were informed that the researchers aimed to use a three-point scale namely "no", "partial" and "yes", which corresponds to 0, 1, and 2 in the rubric, respectively. The directions to score the items in the rubric were provided with the rubric (See AppendixB). To ensure the reliability of the rating, it was checked whether each expert understood the rating procedure properly. They were asked to evaluate the dissertation through the rubric by choosing "no", "partial" and "yes" on a scale in the rubric. Krippendorff's Alpha and Fleiss' Kappa values were calculated to examine the consistency of assessments among the raters in this study. Krippendorff's α was calculated through a macro produced for the SPSS package program by Hayes and Krippendorff (2007).

Table 3. Krippendorff's alpha values for the rubric development in the field of educational sciences

A	95% CI		Units	Raters	Pairs
	Lower	Upper			
.912	.851	.9712	10	3	30

As seen in Table 3, Krippendorff's alpha values for the dissertations implemented in the field of educational sciences were found to be .912. This value is acceptable [α>0.667] (Krippendorf, 2004). Fleiss' Kappa value for these dissertations was calculated to support the former result. According to Table 4, Fleiss' kappa values [κ=.703, p < .05] show a substantial agreement among the raters (Landis & Koch, 1977). The rubric can be regarded as sufficiently acceptable as a result of these calculations.

Table 4. Fleiss kappa values for the rubric development in the field of educational sciences

K	Asymptotic S.E.	z	p	95% Asymptotic CI Bound	
				Lower	Upper
.703	.118	5,951	.000	.471	.935

On the other hand, Krippendorff's alpha values and Fleiss Kappa values were also calculated for the dissertations implemented in the field of social sciences.

Table 5. Krippendorff's alpha values for the rubric development in the field of social sciences

A	95% CI		Units	Raters	Pairs
	Lower	Upper			
.952	.932	.970	10	3	30

As seen in Table 5, Krippendorff's alpha values were found to be .952. This value is acceptable [α>0.667] (Krippendorf, 2004). Fleiss' Kappa value was calculated to support the former result.

Table 6. Fleiss kappa values for the rubric development in the field of social sciences

K	Asymptotic S.E.	z	p	95% Asymptotic CI Bound	
				Lower	Upper
.341	.071	4.819	.000	.202	.480

According to Table 6, Fleiss' kappa values [κ=.703, $p < .05$] show a substantial agreement among the raters (Landis & Koch, 1977). The rubric can be regarded as sufficiently acceptable as a result of these calculations. So, it can be used in the evaluation of Ph.D. dissertations completed in the fields of educational and social sciences.

FUTURE RESEARCH DIRECTIONS

There are several limitations that should be considered dealing with the findings of this study. The current rubric was developed to evaluate Ph.D. dissertations in educational and social sciences. To use the rubric in other disciplines such as medical or engineering sciences, it may be necessary to ensure its validity and reliability. Also, this new rubric was developed mostly based on TQF and TGER descriptors; hence, not taking into account other countries' qualifications frameworks may be a shortcoming of this study. Since comparative studies were not done by experts working in different countries, it is difficult for us to know whether using the same criteria in examining dissertations is appropriate.

CONCLUSION

This study aims to determine the standard criteria for the evaluation processes of Ph.D. dissertations to improve their quality and lessen the differences between the current assessment procedures. Therefore, the new rubric was developed to evaluate the PhD dissertations conducted in the fields of educational and social sciences with regard to the EQF, TQF, TGER, the literature review, and the perceptions of 12 experts in their fields in Türkiye.

To ensure the rubric's reliability, the interrater agreement among the raters in the fields of educational and social sciences was checked separately. To identify the agreement, both Krippendorff's α and Fleiss'

kappa values among the three raters with a Ph.D. degree in educational sciences were calculated. Krippendorff's Alpha was 0.912 and Fleiss' Kappa was 0.703 for the educational dissertations. On the other hand, three raters in the field of social sciences rated 10 social sciences dissertations independently. Krippendorff's Alpha was 0.952 and Fleiss' Kappa was 0.341 for these dissertations. According to the review of 75 studies on rubric conducted by Jonsson and Svingby (2007), it was revealed that many estimates did not reach the criterion of 70% or greater. Stemler (2004) argued that 70% or greater is required for the exact agreement. Two types of interrater reliability methods were employed to enable the reliability of the current rubric. In this regard, it can be said that it met the conditions, which are necessary steps in a rubric development process with regard to reliability. As a result, it is seen that the interrater reliability of the current rubric is acceptable enough.

As in the case of the reliability of the rubric, not only content but also construct validity was implemented by consulting the experts' perceptions. The content validity ratio and content validity index values were calculated to ensure the content validity. These results indicated that it is acceptable with regard to content validity.

The rubric consists of 11 items with three dimensions: theoretical framework, method, and research contributions. There are three items in the theoretical framework, three items in the method, and five items in the contributions of research that describe different criteria for the high quality of dissertations. During the evaluation process of dissertations, these dimensions might be used as a whole or separately in practice.

The rubric appeals to all Ph.D. students, academics, supervisors, faculty, and external jury members. While doctorate students prepare their proposals for their dissertations, they can use it as a guideline. From this perspective, the rubric could have an impact on researchers to produce high-quality dissertations, which have high-value-added, sound methodology and wide impact criteria. In addition, supervisors can guide their students to prepare their dissertations regarding the items in the rubric. Faculty and external jury members can evaluate dissertations based on the common criteria in the rubric. Furthermore, the rubric can indicate to what extent Ph.D. holders are qualified according to each dimension of the rubric. This can provide policymakers and researchers with empirical and sound findings to improve doctoral training in their countries.

The newly developed rubric in this study is expected to enable the doctorate students in the field of educational and social sciences to prepare their Ph.D. dissertations at least the average standards. Hereby, doctorate students can conduct their dissertation studies with regard to the items related to the three determined dimensions: theoretical framework, method, and their studies' contributions to the field. So, the newly developed rubric can promote doctorate students to assess themselves. In addition, it can enable jury members to assess dissertations in the light of the standard criteria rather than their perceived assumptions. Also, this rubric might give us some information about to what extent doctorate students acquire the qualifications defined in TQF and EQF. In sum, the rubric might be used to improve the quality of dissertations. Thus, it might be possible to measure or monitor the quality of dissertations in a meaningful way at an aggregate level. To this end, this study might provide decision-makers with reference to improve the quality of doctorate programs around the world.

DISCUSSION

The rubric can be used as a guideline to enhance educational and social sciences doctoral training programs with respect to curriculum and graduates' competencies based on the concrete outputs of programs, namely Ph.D. dissertations. Academics and policymakers can determine the strengths and weaknesses of the programs with regard to the theoretical framework, method and contributions of research dimensions. To illustrate, when academics, namely jury members observe problems stemming from students' theoretical knowledge in their dissertations, they can make improvements in the curriculum, thereby making them take more theoretical courses in the program. With the improvements made in the programs, it can be possible to train more competent researchers in their study fields.

REFERENCES

Agu, N. N., Omenyi, A. S., & O, C. (2015). Evaluation of doctorate dissertation in Nigerian Universities: do faculties provide and use criteria / rubrics? *International Journal of Technology and Inclusive Education (IJTIE), 4*(1), 565-569.

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, *53*(1), 27–30. doi:10.3200/CTCH.53.1.27-31

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement & Evaluation in Counseling & Development*, *47*(1), 79–86. doi:10.1177/0748175613513808

EQF. (2005). *Descriptors defining levels in the European Qualifications Framework*. EQF. https://ec.europa.eu/ploteus/tr/node/1440

Fitt, M. H., Walker, A. E., & Leary, H. M. (2009). Assessing the quality of doctoral dissertation literature reviews in instructional technology. *Paper presented at the Annual meeting of the American Educational Research Association San Diego*, 3-15.

Hall, C. M. (1991). Tourism as the Subject of Post-graduate Dissertations in Australia. *Annals of Tourism Research*, *18*(3), 520–523. doi:10.1016/0160-7383(91)90061-F

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89. doi:10.1080/19312450709336664

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Karslı, M. B., Karabey, S. Ç., Nergiz, E., & Goktas, Y. (2018). Comparison of the discussion sections of PhD dissertations in educational technology: The case of Turkey and the USA. *Scientometrics*.

Kirk, R. E. (2007). *Statistics, an introduction* (5th ed.). Thomson Wadsworth.

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, *30*(3), 411–433. doi:10.1111/j.1468-2958.2004.tb00738.x

Kyvik, S., & Thune, T. (2015). Assessing the quality of PhD dissertations. A survey of external committee members. *Assessment & Evaluation in Higher Education*, *40*(5), 768–782. doi:10.1080/02602938.2014.956283

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. doi:10.2307/2529310 PMID:843571

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. doi:10.1111/j.1744-6570.1975.tb01393.x

Lovitts, B. E. (2005). How to grade a dissertation. *Academe, 91*(6), 18-23.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, *7*(10), 1–6.

MYK. (2015). *Turkish Qualifications Framework*. MYK. https://www.myk.gov.tr//TRR/File6.pdf

Özen, S. O. (2019). *Öğrenenlerin e-değerlendirmeye dayalı kişiselleştirilmiş geri bildirim yollarının aranması [An investigation of the learners' personalized feedback paths based on e-assessment.]* [Unpublished PhD dissertation, Eskişehir Osmangazi Univeristy, Eskişehir, Turkey].

Parlak, B., & Doğan, N. (2014). Dereceli puanlama anahtarı ve puanlama anahtarından elde edilen puanların uyum düzeyleri [Comparison of answer key and scoring rubric for the evaluation of the student performance]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi / H.U. Journal of Education*, *29*(2), 189–197.

Popham, W. J. (1997). What's wrong-and what's right-with rubrics? *Educational Leadership*, *55*(2), 72–75.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Stellmack, M. A., Konheim-Kalkstein, Y. L., Manor, J. E., Massey, A. R., & Schmitz, J. A. P. (2009). An assessment of reliability and validity of a rubric for grading APA-style introductions. *Teaching of Psychology*, *36*(2), 102–107. doi:10.1080/00986280902739776

Stemler, S. E. (2004). A comparison of consensus, consistency and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, 9.

Texas A&M Commerce University (2016). *Rubrics for assessing dissertations*. Texas A & M Commerce University. https://www.tamuc.edu/academics/graduateschool/Thesis%20and%20Dissertation%20Services/graduate%20rep%20rubrics.aspx

TGER (2016). Turkish Graduate Education Regulation. *27561 numbered Official Gazette*.

Yurdugül, H. (2005). *Ölçek Geliştirme Çalışmalarında Kapsam Geçerliği için Kapsam Geçerlik İndekslerinin Kullanılması.* Paper presented at XIV Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi, Denizli. http://yunus.hacettepe.edu.tr/~yurdugul/3/indir/PamukkaleBildiri.pdf

Zimmaro, D. M. (2004). *Developing grading rubrics*. BSU. http://bsuenglish101.pbworks.com/f/rubricshandout.pdf

ADDITIONAL READINGS

Arı, G. (2021). An overview of writing rubrics in doctoral dissertations in Turkey. *International Journal of Progressive Education*, *17*(2), 385–405. doi:10.29329/ijpe.2021.332.24

Maki, P. L., & Borkowski, N. A. (Eds.). (2006). *The assessment of doctoral education. Emerging criteria and new models for improving outcomes*. Stylus Publishing.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*(0), 129–144. doi:10.1016/j.edurev.2013.01.002

Rakedzon, T., & Baram-Tsabari, A. (2017). To make a long story short: A rubric for assessing graduate students' academic and popular science writing skills. *Assessing Writing*, *32*, 28–42. doi:10.1016/j.asw.2016.12.004

Rivas, M. A., De La Serna, M. C., & Martínez-Figueira, E. (2014). Electronic rubrics to assess competences in ICT subjects. *European Educational Research Journal*, *13*(5), 584–594. doi:10.2304/eerj.2014.13.5.584

The University of North Carolina at Chapel Hill. (2017). *Using rubrics to assess student learning outcomes at the program level.* UNC. https://oira.unc.edu/wp-content/uploads/sites/297/2017/07/Developing-and-Using-Rubrics.pdf

Vaccari, D., & Thangam, S. (2010). *A proposed doctoral assessment procedure and rubric for science and engineering. American Society for Engineering Education* Paper presented at 2010 Annual Conference & Exposition, Louisville, Kentucky. 10.18260/1-2--16106

Yağız, O., Ötügen, R., Kaya, F., & Aydın, B. (2014). A literature review analysis of the Turkish scholars' research articles in ELT and applied linguistics. *Procedia: Social and Behavioral Sciences*, *158*, 389–393. doi:10.1016/j.sbspro.2014.12.105

KEY TERMS AND DEFINITIONS

Dissertation: An essay or thesis written on a particular subject by a candidate to earn the doctorate degree.

Doctor of Philosophy: It refers to the highest degree in an academic level for graduate study.

Educational Sciences: It is any branch of academic study to describe, understand and prescribe the policies and practices to be developed in education.

European Qualification Framework: It is a common European reference framework to make qualifications more readable and understandable across different countries and systems.

Qualifications: They refer to the certificates and diplomas awarded following teaching and learning processes.

Rubric: It is a scoring tool that is used in qualitative rating to evaluate original or complex student works.

Social Sciences: It is any branch of academic study or science that deals with human behaviour in its social and cultural aspects.

Turkish Qualification Framework: It refers to the national qualifications framework which was prepared in line withthe European Qualifications Framework (EQF).

APPENDIX A

Table 7. Graded category rating scale

Dimensions	Item	No	Partial	Yes
THEORATICAL FRAMEWORK	1. The scientific importance of the dissertation was expressed clearly and straightforwardly.			
	2. The research problem of the dissertation was expressed clearly and straightforwardly.			
	3. The theory / theories, on which the dissertation is based, were expressed clearly and straightforwardly.			
METHOD	4. The study group / sample of the dissertation were appropriately determined in line with its aims.			
	5. The validity and reliability treatments of the dissertation were conducted in line with its aims.			
	6. The appropriate methods and techniques in the dissertation were used in line with its aims.			
CONTRIBUTIONS OF RESEARCH	7. The dissertation provided new and authentic contributions to the current literature or applications.			
	8. The dissertation provided new and authentic suggestions to the current literature or applications.			
	9. The dissertation has publication potential.			
	10. The dissertation makes contributions to one of economic, social and cultural fields.			
	11. The dissertation results make associations among different disciplines.			

APPENDIX B

Table 8. Rubric to develop dissertations conducted in educational and social sciences

Dimensions	Item	No	Partial	Yes
THEORATICAL FRAMEWORK	1. The scientific importance of the dissertation was expressed clearly and straightforwardly.	* The scientific importance of the dissertation was not expressed. * Supportive evidences with regard to the scientific importance of the dissertation were not provided.	* The scientific importance of the dissertation was expressed. However, it was not expressed clearly and straightforwardly * Weak evidences with regard to the scientific importance of the dissertation were provided.	* The scientific importance of the dissertation was expressed clearly and straightforwardly. * Strong and persuasive evidences with regard to the scientific importance of the dissertation were provided.
	2. The research problem of the dissertation was expressed clearly and straightforwardly.	* The research problem of the dissertation was not expressed. * The research problem was not supported with evidences. * The groups or cases affected from the research problem were not expressed.	* The research problem of the dissertation was expressed. However, it was not expressed clearly and straightforwardly. * The research problem was supported with weak evidence. * The groups or cases affected from the research problem were ambiguously expressed.	* The research problem of the dissertation was expressed clearly and straightforwardly. * The research problem was supported with strong and empirical evidences. * The groups or cases affected from the research problem were clearly expressed.
	3. The theory / theories, on which the dissertation is based, were expressed clearly and straightforwardly.	* The theory / theories, on which the dissertation is based, were not expressed. * The theory / theories which are not directly related with the dissertation were mostly expressed. * The contradictory or wrong information concerning the theory / theories, on which the dissertation is based, were provided.	* The theory / theories, on which the dissertation is based, were provided. However, it was not expressed clearly and straightforwardly. * The theory / theories, on which the dissertation is based, were not presented in a logical order and holistically. * The contradictory or wrong information concerning the theory / theories, on which the dissertation is based, were provided by mistake.	* The theory / theories, on which the dissertation is based, were expressed clearly and straightforwardly. * The theory / theories, on which the dissertation is based, were presented in a logical order and holistically. * The contradictory or wrong information concerning the theory / theories, on which the dissertation is based, were not provided.
METHOD	4. The study group / sample of the dissertation were appropriately determined in line with its aims.	* The method to determine the study group / sample of the dissertation was not expressed. * The method to determine the study group / sample of the dissertation was not consistent with the study subject and method of the dissertation. * Even though the sample of the study was determined, the method (s) used to determine the sample size was not expressed.	**Only** one point is met: * The method to determine the study group / sample of the dissertation was expressed. * The method to determine the study group / sample of the dissertation was consistent with the study.	* The method to determine the study group / sample of the dissertation was expressed. * The method to determine the study group / sample of the dissertation was consistent with the study subject and method of the dissertation. * The method (s) used to determine the sample size and sample of the study was expressed.
	5. The validity and reliability treatments of the dissertation were conducted in line with its aims.	* Appropriate evidences for the validity of the dissertation were not provided. *Appropriate evidences for the reliability of the data used in the dissertation were not provided. * If a measurement instrument was used in the dissertation; - The validity and reliability treatments of the measurement instrument were not conducted. - If the measurement instrument was developed beforehand, the information concerning its validity and reliability was not reported.	* Weak evidences for the validity of the dissertation were provided. * Weak evidences for the reliability of the data used in the dissertation were provided. * If a measurement instrument was used in the dissertation, its validity and reliability treatments were conducted. However, they did not meet the expectations.	* Appropriate and strong evidences for the validity of the dissertation were provided. * Appropriate and strong evidences for the reliability of the data used in the dissertation were provided. * If a measurement instrument was used in the dissertation; - Its validity and reliability treatments were appropriately conducted. - If the measurement instrument was developed beforehand, the information concerning its validity and reliability was reported.
	6. The appropriate methods and techniques in the dissertation were used in line with its aims.	* The inappropriate methods and techniques were used in the dissertation. * The data collection instrument which was not appropriate for the dissertation topic and method was used.	* The relative appropriate methods and techniques were used in the dissertation. * The data collection instrument which was relatively appropriate for the dissertation topic and method was used.	* The appropriate methods and techniques were used in the dissertation. * The data collection instrument which was appropriate for the dissertation topic and method was used.
	7. The dissertation provided new and authentic contributions to the current literature or applications.	* The dissertation results did not provide new and authentic contributions to the current literature or applications. * The results of the similar studies were repeated.	* The dissertation results provided new and but relatively authentic contributions to the current literature or applications. * The results of the similar studies were partly extended.	* The dissertation results provided new and authentic contributions to the current literature or applications. * The scope of the literature or applications was extended and different perspectives were provided.

Continued on following page

Table 8. Continued

Dimensions	Item	No	Partial	Yes
	8. The dissertation provided new and authentic suggestions to the current literature or applications.	* New and authentic suggestions were not provided to the current literature or applications based on the dissertation findings. * The suggestions made in the similar studies were provided. * The suggestions which were not based on the dissertation findings were developed.	* New and but relatively authentic suggestions were provided to the current literature or applications based on the dissertation findings. * The suggestions provided in similar studies were partly repeated.	* New and authentic suggestions were provided to the current literature or applications based on the dissertation findings. * The scope of the suggestions made in other studies was extended. * All the suggestions provided in the dissertation were consistent with its findings.
	9. The dissertation has publication potential.	* The dissertation does not have potential to be published in moderate or high impact indexed journals*. *ESCI, SSCI, AHCI, SCI, SCI Expanded, ERIC, SCOPUS indexed journals.	*The dissertation has potential to be published in moderate impact indexed journals*. *ESCI, SCOPUS, ERIC indexed journals.	* The dissertation has potential to be published in high impact indexed journals*. *SSCI, AHCI, SCI, SCI Expanded indexed journals.
CONTRIBUTIONS OF RESEARCH	10. The dissertation makes contributions to one of economic, social and cultural fields.	*Any contributions were not made to economic, social and cultural fields.	* A new but relatively authentic approach was contributed to one of economic, social and cultural fields.	A new and authentic approach was contributed to one of economic, social and cultural fields.
	11. The dissertation results make associations among different disciplines.	* The dissertation results did not make associations among different disciplines.	* The dissertation results made narrow scoped associations among different disciplines.	* The dissertation results made profound and comprehensive associations among different disciplines.