New Perspectives in Science Education
International Conference
NEW PERSPECTIVES in SCIENCE EDUCATION

CoRubric as a Tool to Improve Argumentation by Peer-Assessment in Pedagogy Students

Daniel Cebrián-Robles[1], Antonio-Joaquín Franco-Mariscal[2]

Abstract

Argumentation is considered to be one of the fundamental pillars of science. As such, numerous efforts have been made to evaluate the quality of arguments by way of European programs such as PISA or to analyse argumentation in primary or secondary school preservice teachers. Many of these studies have used rubrics to evaluate arguments. In light of this, herein we present CoRubric as a tool to help in this evaluation process. This research is based on a study with 38 students taking the "Educational evaluation of learning" module in the 3rd year of the Pedagogy Degree at the University of Malaga (Spain). An argumentative task based on the problem of the appearance of tooth decay was proposed as a pilot experience in a socio-scientific framework. The activity required students to argue whether there is a relationship between bacteria and the appearance of caries. Students were told to create arguments and evaluate their partners' contributions in a peer-assessment manner. As a result of this work, an argument evaluation methodology is presented using the free online tool CoRubric. Furthermore, the initial conceptions as to what students understand by arguing in science were analysed and categorized on the basis of what pedagogy students think about how they should assess a scientific argument made by their peers.

Keywords: *argumentation, assessment, pedagogy, e-rubric, CoRubric*

1. Introduction

Argumentation is considered to be one of the main pillars of science teaching [1,2], and the Toulmin model [3] the most suitable model for explaining the structure of an argument. However, as teaching and recognising the elements in an argument can be a complex task, some authors have simplified this model by including only the essential elements [4], namely a conclusion, which aims to provide a statement regarding a controversial topic, evidence/s to support that conclusion and a justification to link this evidence to the conclusion.

Argumentation is also relevant socially as people need to know how to reason in different situations they may come across. As such, it is important to provide future teachers with an understanding of an argumentation model that they can pass on to their students. Unfortunately, the literature shows that trainee teachers and educators, and even serving teachers, are rarely trained in argumentation programmes, a deficiency that means they do not teach how to reason [5]. Despite this, numerous efforts have been made to evaluate the quality of the reasoning used by students. These include European programmes such as the Programme for International Student Assessment (PISA) [6] or research concerning argumentation-based training programmes with students from all levels, from preservice [7,8] to secondary [8,9] or primary education [10].

Although the evaluation of reasoning may be a useful tool in the teaching/learning process [11], this is not a simple task, and much less so if the responsibility for evaluating the quality of reasoning is passed to the students themselves [8]. Despite its importance, very little effort has been dedicated to ensuring that trainee teachers learn to evaluate.

Various instruments can be used to accurately and objectively evaluate the quality of an argument. Although rubrics are the most widely used such instrument [12–14], some authors prefer to use questionnaires [7] or argumentation questions to aid this process [15]. Despite this, peer evaluation is perhaps the most appropriate instrument for initiating students in evaluation training provided it is carried out in accordance with rules that are clear and well-defined for the students [16]. Recent studies concerning peer evaluation when learning to reason have shown an improvement in the quality of the arguments put forward by students [17]. E-rubrics, which allow the results of the evaluations to be shown dynamically and instantaneously, are typically used for this type of evaluation. This study uses CoRubric [18], a tool that aids this peer-evaluation process and allows a range of data concerning the evaluations to be collected, the results to be shown instantaneously, groups to be

[1] University of Malaga, Spain
[2] University of Malaga, Spain

created, evaluator roles to be assigned, data to be exported for analysis, 360-degree evaluation methodologies for self-evaluation, peer co-evaluation and evaluation of the lecturer.
The aim of this study is to demonstrate the pre-fixed ideas held by preservice elementary pedagogy teachers (hereinafter PEPTs) concerning evaluation of the quality of an argument and how peer evaluations slowly converge with those of the lecturer as an argumentative model is explained and the evaluation rubrics are presented.

2. Method

The participants in this study were 38 PEPTs taking the “Educational evaluation of learning” module in the 3rd year of the Pedagogy Degree at the University of Malaga (Spain). This subject includes a module concerning evaluation methodologies and technological tools The study was carried out as part of that module, including a 12-hour training programme in argumentation that explained how to present a good scientific argument and how this can be evaluated. All the tasks proposed as part of that programme had a level equivalent to that of a student upon completing compulsory secondary education as the participating PEPTs had not taken any science subjects or had any contact with them since grade 10 (16 years of age). This paper presents the results corresponding to a task in the programme that was conducted over two sessions.
Session 1: PEPTs were asked to answer an argumentative question drawn from PISA [15] on tooth decay, which addressed the influence of bacteria on the formation of tooth decay. The text of the question was: “Bacteria that live in our mouths cause dental caries (tooth decay). Caries has been a problem since the 1700s when sugar became available from the expanding sugar cane industry. Today, we know a lot about caries. For example: a) Bacteria that cause caries feed on sugar. b) The sugar is transformed to acid. c) Acid damages the surface of teeth. d) Brushing teeth helps to prevent caries (figure 1). Do bacteria play an important role in the onset of dental caries? Justify your response”.

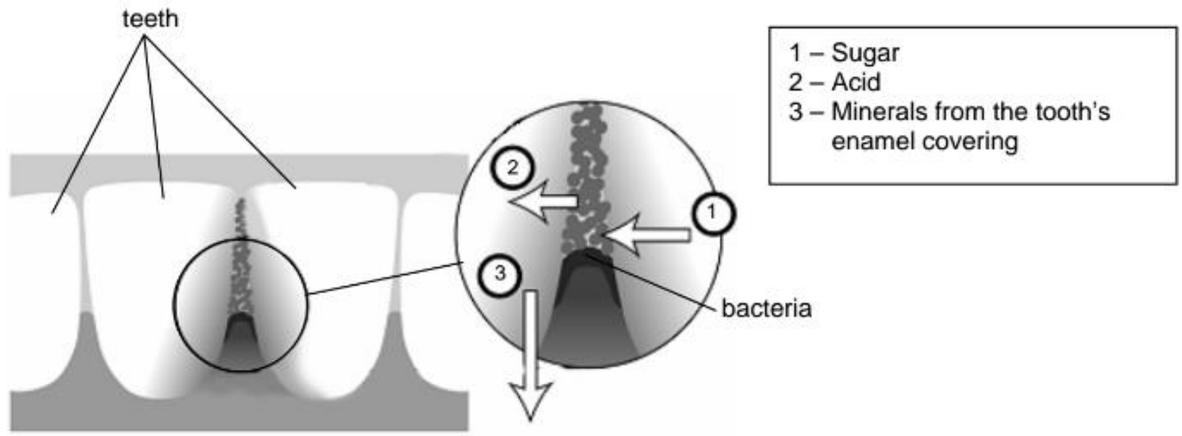

Fig. 1. Image in the tooth decay task [15]

Once the task had been completed, each PEPT was assigned an identifier and asked to evaluate and justify whether the responses given by another two randomly selected PEPTs were well reasoned and score them on a scale of 1 to 10. In parallel, the lecturer, who is also the first author on this paper, evaluated and scored the responses of all PEPTs. The evaluations performed by the PEPTs were coded reflexively and iteratively by the members of the research team, who derived emergent themes (or categories) on the basis of the criteria used [3].
Session 2: The lecturer explained Toulmin’s argumentation model [3] and presented a basic rubric (figure 2) for evaluating the quality of an argument using three element (evidence, justification and conclusion). This basic rubric consisted of a generic rubric that can be adapted to argumentation tasks.

1.Evidence			
1 There is no evidence	2 There is some evidence, but inappropriate to support the conclusion	3 There is some evidence but not sufficient. Some evidence given might be inappropriate	4 There is enough evidence, and it is appropriate to support the conclusion

2.Justification			
1 There is no justification	2 There is a justification, but it does not link the evidence with the conclusion	3 There is a justification that links the evidence with the conclusion. It repeats the evidence and/or some scientific ideas are included but they are insufficient	4 There is a justification that links the evidence with the conclusion. It includes the appropriate and sufficient scientific ideas

3.Conclusion/s			
1 There is/are no conclusions	2 The conclusion is not right	3 The conclusion is right, but it bears scientific mistakes	4 The conclusion is right and bears no scientific mistakes

Fig. 2. General basic rubric to assess arguments

The specific rubric on tooth decay was designed with a different number of achievement levels from the basic rubric in order to reflect whether the question is asked in a tentative way.

1.Evidences				
1 There is no evidence	2 Although there is some evidence, it has nothing to do with bacteria and caries	3 Some evidence that justifies the connection between bacteria and caries is provided but it is insufficient	4 Sufficient and appropriate evidence is provided to establish a link between bacteria and caries	

2.Justifications				
1 There is no justification	2 There is a justification of the conclusion but it does not link the evidence to the conclusion	3 Provided a justification linking the evidence with the conclusion but it is insufficient to argue that there is a link between bacteria and caries	4 Provides a justification linking the evidence with the conclusion. Sufficient and appropriate scientific ideas are presented, the justification lacks conviction	5 Provides a justification linking the evidence with the conclusion. Sufficient and appropriate scientific ideas are presented. In addition, the terms are used accurately and taxatively.

3.Conclusion/s				
1 There is no conclusion (not known or cannot be known)	2 An erroneous conclusion concerning the role of bacteria in caries is reached	3 Although the answer is correct, it is not expressed scientifically	4 The correct conclusion concerning the role of bacteria in caries is reached. However, although expressed scientifically, the conclusion lacks conviction ("I think", "In my opinion", etc.)	5 The correct conclusion concerning the role of bacteria in caries is reached. In addition, it is well expressed scientifically and taxatively

Fig. 3. Assessment rubric for the task on tooth decay

Once the simplified Toulmin model and the task-specific rubric had been explained, the PEPTs re-evaluated the responses of two randomly chosen peers using the CoRubric platform.

To determine the agreement between the scores awarded by the PEPTs in the two sessions and those awarded by the lecturer, the Pearson's correlation coefficient *r* was calculated using the statistical software package SPSS 23.0.

3. Results

3.1. Criteria used by the PEPTs to evaluate the arguments

An analysis of the evaluations by the PEPTs in session 1 allowed their justifications to be categorised on the basis of four criteria, which are ordered on the basis of the percentages obtained (figure 4): a) Justifications based on the argumentation itself (e.g., "*uses information from the lecture*" or "*lack of argumentation*"; b) based on an understanding of the task required and its solution (e.g. "*has correctly interpreted the information provided in the lecture*" or "*overly simple response*"; c) use of language when expressing him-/herself (e.g. "*appropriate vocabulary*" or "*drafting errors*"); d) scientific understanding (e.g. "*clear concepts*" or "*relates concepts*"); e) other aspects (e.g. "*response with a more personal type of reasoning missing*").

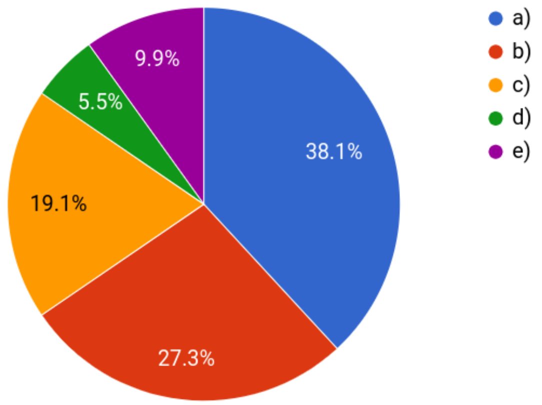

Fig. 4. Percentages in each category for the criteria used by the PEPTs

It can be seen that, as the PEPTs have no prior training in argumentation, they initially evaluated the responses on the basis of their own preconceptions concerning the quality of an argument. Only 38% of the criteria applied in the evaluations are related to argumentation (category a). Scientific understanding (category d) also presents a very low percentage (5.5%). Both these criteria (a and d) are most directly related to the quality of an argument.

3.2. Correlation with the score awarded by the lecturer in the two tasks

The statistical analysis between the scores awarded by the PEPTs and the lecturer in the first evaluation without rubric for session 1 showed no correlation between groups ($r = 0.18$, $p > 0.05$). However, the second time that it was evaluated using the rubric shown in figure 3, this score was more similar to that given by the lecturer, with a mean of Pearson's correlation of $r = 0.30$ ($p = 0.00$) being obtained overall. The following correlations were obtained for each element in an argument: $r = 0.33$ ($p < 0.01$) for evidence and $r = 0.10$ ($p > 0.05$) for justification and conclusions.

4. Conclusions

Our findings show the importance of reaching how to argue and to evaluate an argument [19] as the initial criteria used by the PEPTs are not related to important aspects concerning the quality of arguments.

An understanding of argumentation and the components of an argument improves the quality of the evaluations provided by PEPTs to some degree as they become more similar to those of the lecturer. The CoRubric tool used for peer review most likely also has an influence in this regard. However, the results obtained show the need to offer PEPTs more opportunities to acquire skills concerning argumentation and its evaluation.

Acknowledgments

This work is part of the ‘I+D Excelencia’ Project EDU2017-82197-P funded by the Spanish Ministry of Economy and Finance through its 2017 research call and of the Innovative Education Project PIE 15-161 of the University of Malaga.

References

[1] Duschl, R.A., Osborne, J. "Supporting and Promoting Argumentation Discourse in Science Education", Studies in Science Education, 2002, 38, 39–72.
[2] Taylor, C.A. "Defining Science: A Rhetoric of Demarcation", Wisconsin, The University of Wisconsin Press, 1996.
[3] Toulmin, SE. "The uses of argument", 2003, Cambridge, Cambridge University Press, 1958.
[4] Jiménez-Aleixandre, M.P. "10 Ideas Clave. Competencias en argumentación y uso de pruebas", Barcelona, Graó, 2010.
[5] McNeil, K.L., Knight, A.M. "Teachers’ Pedagogical Content Knowledge of Scientific Argumentation: The Impact of Professional Development on K–12 Teachers", Science Education, 2013, 97, 936–72.
[6] OECD. "The Organisation for Economic Co-operation and Development", 2016.
[7] Aydeniz, M., Ozdilek, Z. "Assessing and enhancing pre-service science teachers’ self-efficacy to teach science through argumentation: Challenges and possible solutions", International Journal of Science and Mathematics Education, 2015, 1–19.
[8] Larson, A.A., Britt, M.A., Kurby, C.A. "Improving Students’ Evaluation of Informal Arguments", Journal of Experimental Education, 2009, 77, 339–65.
[9] Deng, Y, Wang, H. "Research on evaluation of Chinese students’ competence in written scientific argumentation in the context of chemistry", Chemical Education Research and Practice, 2017, 18, 127–150.
[10] Bulgren, J, Ellis, J, Marquis, J. "The Use and Effectiveness of an Argumentation and Evaluation Intervention in Science Classes", Journal of Science Education and Technology, 2014, 23, 82-97.
[11] Folkes, C., Carmichael, P. "“Learning to assess” and “assessing to learn”: the construction of knowledge about Assistive Technology", Educational Action Research, 2006, 14, 535-45.
[12] Lu, J., Zhang, Z. "Assessing and Supporting Argumentation with Online Rubrics", International Education Studies, Toronto, 2013, 6, 66.
[13] Knight-Bardsley, A., McNeill, K.L. "Teachers’ Pedagogical Design Capacity for Scientific Argumentation", Science Education, 2016, 100, 645–72.
[14] Jonsson, A., Svingby, G. "The use of scoring rubrics: Reliability, validity and educational consequences", Educational Research Review, 2007, 2, 130-44.
[15] OECD. "Assessing scientific, reading and mathematical literacy: A framework for PISA 2006", Brussels, OECD, 2006.
[16] Boud, D., Cohen, R., Sampson, J. "Peer Learning and Assessment", Assessment & Evaluation in Higher Education, 1999, 24, 413–26.
[17] Clark, D.B., Sampson, V. "Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality", Journal of Research in Science Teaching, 2008, 45, 293-321.
[18] Cebrián-Robles, D. "CoRubric", 2016.
[19] Osborne, J.F., Henderson, J.B., MacPherson, A., Szu, E., Wild, A., Yao, S. "The development and validation of a learning progression for argumentation in science", Journal of Research in Science Teaching, 2016, 53, 821–46.