ESERA 2017 Conference
Dublin City University, Dublin, Ireland.
21st - 25th August 2017

CAPACITY OF EVALUATION OF PRESERVICE ELEMENTARY SCIENCE TEACHERS IN AN ARGUMENTATION TASK

Daniel Cebrián-Robles[1], Antonio Joaquín Franco-Mariscal[1] Ángel Blanco-López[1]
[1] Universidad de Málaga, Didáctica de las Ciencias Experimentales, Málaga, Spain

Abstract: The growing importance of argumentation in science education must also be accompanied by examples showing both teachers in service and trainee-teachers how to implement and assess argumentation in class. In this line, this study is framed within a broader research study on argumentation competency , which involves the participation of Preservice Elementary Science Teachers (PEST)s from 3rd year of the Primary Education Teaching Degree from the University of Malaga (Malaga, Spain). Specifically, this paper shows an argumentation task that involves the participation of 98 Spanish Preservice Elementary Science Teachers (PESTs), through production and peer assessment, in order for them to internalise the criteria of a good argument, thus improving their argumentation skills. The set task is drawn from a PISA 2006 test, which addresses the possibility of reducing the hardness of a lipstick by changing its composition. PESTs are required to conduct peer assessment, which will then be compared to teacher assessment, in order to analyse the ability of the former to identify and evaluate the elements of an argument. Results show different levels of capacity for analysis and evaluation of argumentation by PESTs. In particular, they struggle the most to identify and evaluate evidence and justification. Likewise, the peer assessment-teacher assessment comparison reveals an overestimation by PESTs in relation to evidence and justification.

Keywords: Argumentation, Preservice Elementary Science Teachers, Peer Assessment.

INTRODUCTION

Nowadays, argumentation is considered one of the main scientific practices, thus a key element in science teaching (Erduran & Jiménez-Aleixandre, 2008). McNeil & Knight (2013) consider it necessary to explicitly work out with future trainee-teachers and teachers in service the best way to address argumentation in class. In order to design and evaluate argumentative activities, an appropriate model for understanding argumentation is needed. With regards to this, Toulmin's model has been simplified by Jiménez-Aleixandre (2010) to facilitate the understanding of the essential elements that a good argument must have: evidence (E), justification (J) and conclusions (C). Different authors agree that, in order to understand and internalise the criteria for a good argument, one needs to practice such criteria, by producing arguments and assessing those of others (Osborne et al., 2016). A number of research studies prove that students learn to assess when they actually assess to learn (Boud, Cohen & Sampson, 1999). Likewise, other studies highlight that teamwork, where students can see and assess their peers' arguments, allows them to improve the quality of their arguments (Evagorou & Osborne, 2013). Some authors use questionnaires to assess the elements of an argument, (Clark & Sampson, 2008), while others prefer rubrics (Deng & Wang, 2017). Considering the above ideas, this study focuses on analysing the difficulties encountered by PESTs to assess peers' arguments, and on comparing these assessments to those conducted by teachers.

METHOD

This study is framed within a broader research study on argumentation competency (Osborne et al., 2016), which involves the participation of 98 PESTs from 3rd year of the Primary Education Teaching Degree from the University of Malaga (Malaga, Spain). Students belong to two different, randomly chosen, groups. The study shows the results of one of the tasks performed during a training programme aimed at teaching students to argue. More specifically, the task is performed after introducing Toulmin's model and the

elements of a good argument, plus presenting several examples of tasks on argumentation and evaluation. The set task poses PESTs the possibility of reducing the hardness of a lipstick by changing its composition. This task was adapted from a science test by PISA 2006, which considers that "the context of cosmetics has everyday relevance for students of this age group, although it could be expected that this task would generate more interest among females than males" (OECD, 2006, p.153). A specific rubric (Fig. 1) was designed for peer assessment and teacher assessment of PESTs. The rubric shows different response levels for each element of an argument, based on a 1 to 4-5 scale, where level 4 for E, level 5 for J and level 5 for C are most desirable.

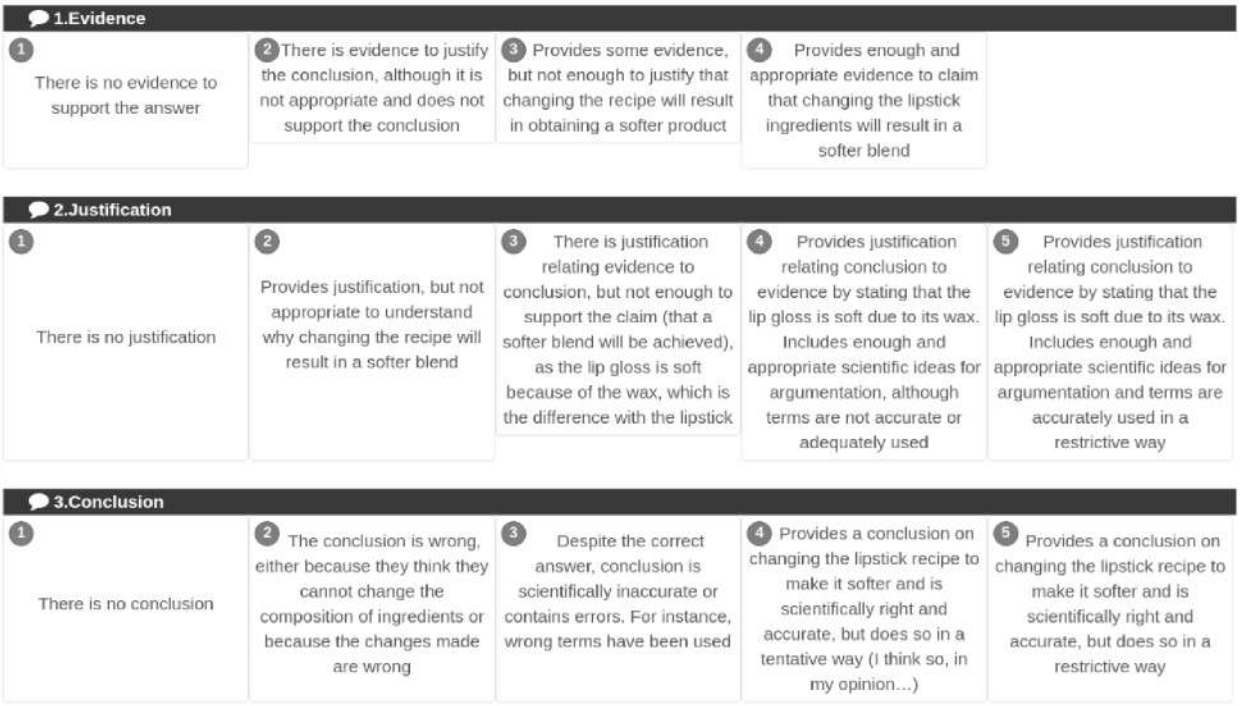

Figure 1. Specific rubric to assess the task on "hardness of a lipstick"

Once the task has been answered, PESTs have to anonymously and randomly assess the responses of two other peers, through the CoRubric electronic rubric collaborative platform (Cebrián-Robles, 2016). The assessment involves identifying three elements in arguments (E, J and C) in their peers' answers and use rubrics to assess the quality of each argument. Similarly, teachers participating in the training programme (two of the authors of this paper) are required to assess PESTs' answers. To do so, they are to agree on the scores assigned to each answer.

RESULTS

Table 1 shows PESTs rates in each achievement level of E, J and C; based on peer assessment and teacher assessment of PESTs.

Table 1. Response rates per achievement level for E, J and C; assessed by PESTs and teachers.

	Peer Assessment (PESTs)						Teacher Assessment					
	Achievement Level (%)					Mean	Achievement Level (%)					Mean
	5	4	3	2	1		5	4	3	2	1	
Evidence (E)	---	40.5	40.5	11.8	7.3	3.14/4	---	29.9	35.1	6.2	28.9	2.66/4
Justifications (J)	22.3	34.1	26.4	12.7	4.5	3.57/5	15.3	20.4	17.3	20.4	26.5	2.77/5
Conclusions (C)	39.1	33.6	21.8	4.1	1.4	4.05/5	68.4	20.4	7.1	4.1	0.0	4.53/5
Total	20.5	36.1	29.5	9.5	4.4		28.0	23.5	19.8	10.2	18.4	

PESTs assess E and C responses in the two highest levels (81% of E in levels 3-4 and 72% of C in levels 4-5). As for J, the response rate concentrates at low levels (60.5% in levels 3-4). Rates found in the lowest level for E, J and C do not exceed 7.3% in any case. These results differ from teachers' assessment, who consider fewer response rates in E and J to be in the highest levels (10% and 7% less, respectively), and 30% more responses in line with a very appropriate conclusion. Finally, teachers assess higher response rates at levels 1-2 in all cases except for C.

CONCLUSIONS AND PROPOSALS FOR IMPROVEMENT

This research study shows a type of argumentation in class that aims to bring PESTs closer to the socio-scientific contexts of their daily lives. The tasks can help increase their level of motivation for science while improving their argumentative competence (Osborne et al., 2016), not only by producing arguments but also by assessing and identifying the main elements of a good argument through peer assessment. The aforementioned results reveal that PESTs struggle to identify and assess E and J, which are often overvalued in relation to teacher assessment. Likewise, some PESTs undervalue C, by a difference of up to two levels. This could be due to PESTs' difficulty to clearly differentiate the three elements in an argument. Results suggest the need to train PESTs in argumentation tasks, with special emphasis on the meaning and use of E and J in arguments. The strategy of peer assessment used in this study seems to have been useful to improve PESTs' argumentation skills, as not only it makes them participate in assessment but also makes them aware of their own argumentative level, which enables them to internalise, be critical and reflect about their own arguments as well as those of others. All of which helps improve self-regulated learning and the argumentative competence.

ACKNOWLEDGEMENT

This work is part of the 'I+D Excelencia' Project EDU2013-41952-P funded by the Spanish Ministry of Economy and Finance through its 2013 research call.

REFERENCES

Boud, D., Cohen, R., & Sampson, J. (1999). Peer Learning and Assessment. *Assessment & Evaluation in Higher Education,* 24(4), 413-426.

Cebrián-Robles, D. (2016). CoRubric. Retrieved from: http://corubric.com

Clark, D.B., & Sampson, V. (2008). Assessing Dialogic Argumentation in Online Environments to Relate Structure, Grounds and Conceptual Quality. *Journal of Research in Science Teaching,* 45(3), 293-321.

Deng, Y., & Wang, H. (2017). Research on Evaluating Chinese Students' Competence of Written Argumentation in the Context of Chemistry. *Chemical Education Research & Practice,* 18, 127-150.

Erduran, S., & Jiménez-Aleixandre, M. P. (2008). *Argumentation in Science Education*. Berlin: Springer.

Evagorou, M., & Osborne, J. (2º013). Exploring Young Students' Collaborative Argumentation within a Socio-Scientific Issue. *Journal of Research in Science Teaching,* 50(2), 209-237.

Jiménez-Aleixandre, M.P. (2010). *10 ideas clave: competencias en argumentación y uso de pruebas* [10 Key Ideas: Argumentation Skills and Use of Evidence] Barcelona: Graó.

McNeil, K.L., & Knight, A.M. (2013). Teachers' Pedagogical Content Knowledge of Scientific Argumentation: The Impact of Professional Development on K–12 Teachers. *Science Education*, 97(6), 936-972.

OECD (2006). PISA. Assessing Scientific, Reading and Mathematical Literacy: *A Framework for PISA 2006. Annex A, Additional Science Units*. Unit 9. Lip Gloss, pp. 153-155. Brussels: OECD.

Osborne, J.F., Henderson, J.B., MacPherson, A., Szu, E., Wild, A. & Yao, S. (2016). The Development and Validation of a Learning Progression for Argumentation in Science.*Journal Research in Science Teaching,* 53(6), 821-846.