PROPOSAL FOR A FRAMEWORK TO EVALUATE ELEMENTARY SCHOOL STUDENTS UNDERSTANDING OF NATURAL SELECTION

Xana Sá-Pinto
Centro de Investigação em Didática e Tecnologia na Formação de Formadores, Universidade de Aveiro; P.Porto: Escola Superior de Educação.

Alexandre Pinto
P.Porto: Escola Superior de Educação; Centro de Investigação em Didática e Tecnologia na Formação de Formadores, Universidade de Aveiro.

Pedro Cardia
P.Porto: Escola Superior de Educação.

Maria João Fonseca
CIBIO-InBIO – Centro de Investigação em Biodiversidade e Recursos Genéticos, Rede de Investigação em Biodiversidade e Biologia Evolutiva, Laboratório Associado; MHNC-UP – Museu de História Natural e da Ciência da Universidade do Porto

Joaquim Bernardino Lopes
Universidade de Trás-os-Montes e Alto Douro; Centro de Investigação em Didática e Tecnologia na Formação de Formadores.

SUMMARY: Evolution is one of the key concepts in biology that should be explored since kindergarten. However, the few studies analysing elementary school students' learning about natural selection used distinct criteria to evaluate their knowledge. In the present work we develope a framework to evaluate students' understanding of natural selection, based on a literature review and on an empirical study. This framework can be used to assess students understanding of evolution and natural selection, to inform the development of educational activities and to assess their impact on students' understanding of these processes.

KEY WORDS: Evolution, natural selection, evaluation framework, elementary school students.

OBJECTIVES: This work seeks to provide a framework to evaluate students' understanding of evolution by natural selection. This evaluation framework is essential to study what students can learn about this process at different developmental stages, to set learning progression goals, to inform the design of educational activities tailored to address students' needs in a specific class and to help develop, test and compare impacts of pedagogical interventions designed to foster evolution understanding.

THEORETICAL FRAMEWORK

Evolution allows us to understand and predict several aspects of the world, as all species and ecosystems result from evolutionary processes acting on their features and ecological interactions. In an educational context, evolution 1) allows students to articulate concepts from distinct disciplines and

integrate them within a wider framework (National Research Council [NRC], 2012), 2) promotes a clear understanding of topics in biology and 3) constitutes a conceptual tool that students can use to explore and tackle new problems (Jenkins, 2009). NRC (2012) has recently proposed evolution as one of the four core ideas around which learning progressions in biological sciences should be built since kindergarten. In agreement with this view, researchers as well as scientific and educational organisations have been highlighting the importance of exploring evolution and evolutionary processes from an early stage (Nadelson *et al.*, 2009; Wagler, 2010; NRC, 2012; Campos & Sá-Pinto, 2013).

Despite of this, and up to our knowledge, only few studies have so far analysed elementary school students' understanding of natural selection (Campos & Sá-Pinto, 2013; Kelemen, Emmons, Schillaci & Ganea, 2014; Berti, Barbetta & Toneatti, 2015). These showed that pedagogical interventions contribute to increase elementary students understanding of evolution, but differences were found in the degree of students understanding of natural selection. Campos & Sá-Pinto (2013) and Kelemen, Emmons, Schillaci & Ganea (2014) reported that, after distinct pedagogic interventions, elementary school students were able to understand and apply the principle of natural selection to explain and predict biological evolution. However, in a study testing a distinct pedagogical sequence, Berti, Barbetta & Toneatti (2015) reported that only a minority of children were able to learn about natural selection. These contradicting results may be explained by differences in the pedagogical interventions and learning contexts (Berti, Barbetta & Toneatti, 2015), but can also stem from differences in the procedure employed by these authors to evaluate natural selection understanding. In fact, very distinct approaches were used to evaluate students' understanding of natural selection. This highlights the importance of finding criteria that can be used to evaluate students' answers and to clearly characterise their conceptions and understanding of evolution by natural selection and the impact of educational interventions.

METHODOLOGY

To analyse natural selection understanding, a convenience sample composed by two K3 (8-9 years old) and three K4 classes (9-10 years old) were recruited from one elementary school located in the most populated civil parish of a municipality in Viseu's district (central Portugal).

It is important to note that evolution and natural selection are not mentioned in Portuguese elementary school programs (Ministério da Educação 2004) and, according to the teachers were not explicitely explored by the sampled classes.

The test used to assess students understanding of evolution and natural selection presented them a scenario of an insular polymorphic population of birds most of which had small bills (less fit phenotype) with only one individual having a large bill (the fittest phenotype; test available at goo.gl/STSsNZ). Like in Campos & Sá-Pinto (2013) the students were asked to think forward in time and predict the outcome of this scenario by describing how the population would look like in 100 years. This evaluation procedure differs from the one of Kelemen, Emmons, Schillaci & Ganea (2014) who asked students to think backwards in time and already presents an evolutionary scenario, precluding students to provide fixist explanations. The test was read loud to diminish the impact of reading abilities on test performance. Students were asked to draw their predictions and write down its justification. After finishing these tasks, and before delivering the test, each student was individually asked to verbally explain her/his predictions and respective justifications. When students verbally provided more information to the researcher than that written or drawn in the test, they were asked to complete their answer in the test. This procedure was followed independently of the students' type of answer. Between 45 and 60 minutes were necessary in each class to get all students' answers.

The evaluation framework was constructed based on published literature (Campos & Sá-Pinto, 2013; Kelemen, Emmons, Schillaci & Ganea, 2014) and further complemented derived from the analyses of students tests (complete definition of each coding rubric item available at goo.gl/DnUSyi).

Three researchers evaluated all student´s answers: one science education researcher (AP), two evolutionary biologists (XSP and PC) one of which with a background in science education (XSP). Given the evaluators trainning, interrater reliability was estimated as the percentage of initial agreement between evaluaters (McHugh, 2012). Answers not equally rated by the three researchers were discussed and, failling a consensus, removed from the analyses. Spearmans' correlation coefficient and its corresponding statistical significance was estimated to test for correlations between coding rubric items. This procedure was essencial to determine how to rate student´s predictions and to build the evolution understanding evaluation framework that resulted from and is proposed in the present work. To test the utility of the evaluation framework to detect significant differences between classes, Mann-Whitney U and Kruskal-Wallis tests were used for pairwise and multiple case comparisons, respectively. All statistical analyses were performed with SPSSv23.

RESULTS

Interrater reliability was higher than 93% for all ítems analysed. Figure 1 shows the relative frequency of answers that were assigned to a given coding rubric item (examples of answers classified as belonging to a given rubric can be found at the document available at goo.gl/w48v6g).

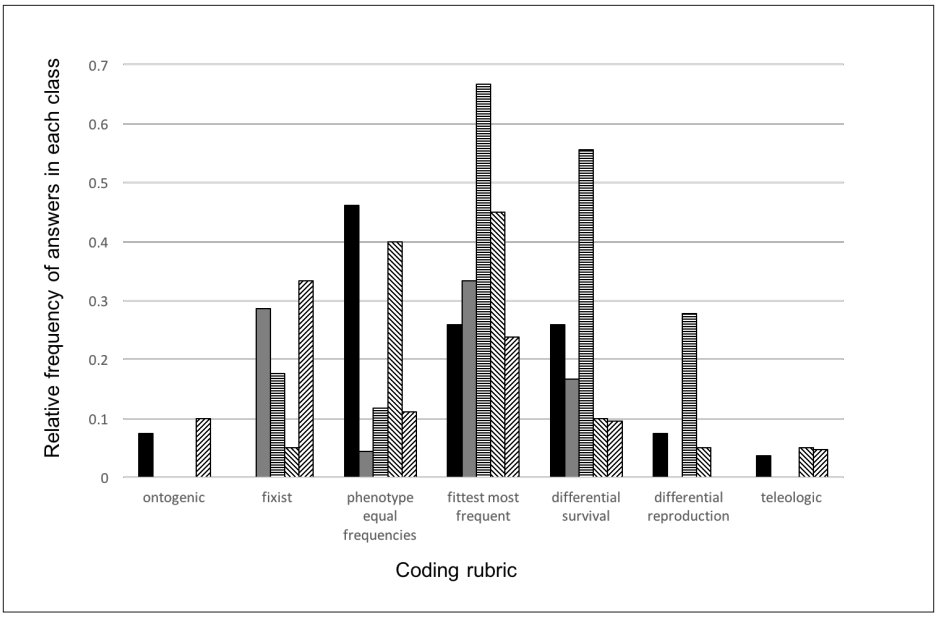

Fig. 1. Relative frequency of answers (Y axis) that could be assigned to a given coding rubric item per class (X axis). Black and grey solid bars represent each of the two K3 classes and streaky bars represent each one of the three K4 classes

Students' most commonly predicted (37.2%) "the fittest will be the most frequent". Nearly 22.7% of the students justified this prediction with differential survival of individuals with distinct bill phenotypes and 7.2% with their differential reproduction. Two students justified their prediction with a teleologic explanation. A positive and significant correlation was found between the prediction "the

fittest will be the most frequent" and justifications mentioning "differential survival" ($r(108)=0.704$, $p<0.01$) and "differential reproduction" ($r(108)=0.363$, $p<0.01$).

Nearly 24% of the students predicted both bill phenotypes would become equally frequent. Whereas this prediction admits population evolution, the justifications offered did not reveal understanding of natural selection. In fact, a negative correlation was found between this prediction and justifications mentioning "differential survival" (r (102)=-0.316, $p<0.01$) and "differential reproduction" (r (102)=-0.162, $p=0.10$). This suggests that despite their evolutionary prediction, these students fail to understand the process of natural selection and these answers were assigned to a distinct subcategory and to a lower level of understanding than the preditction "the fittest will be the most frequent" (Table 1).

Nearly 16% of the students predicted no changes in bill size frequencies (fixist answers), justifying with the initially higher frequency of small billed birds.

Approximately, 3.7% of the students mistook bill size polymorphism for different developmental stages stating in their answer that older birds had larger bill size and younger birds were small billed.

Teleologic explanations were not frequent (1.8%) in our results.

According to these results we propose to organize our coding rubric items in the categories and subcategories depicted in Table 1.

Table 1.
System of categorization (categories, sub-categories and coding rubric items and respective tree organisation). Complete definition of each coding rubric item available at goo.gl/DnUSyi

Categories	Sub-categories	Coding rubric item
Evolutionary	Understanding of Natural Selection	The fittest will be the most frequent
		Differential survival
		Differential reproduction
	Not requiring understanding of Natural Selection	Phenotype equal frequencies
		Teleological explanation
Non evolutionary	Fixist answer	Fixist
	Ontogenic answer	Ontogenic

We further propose to evaluate students' answers according to five levels of evolution understanding: Level 0 (L0): non evolutionary prediction; Level 1 (L1): evolutionary prediction lacking a clear understanding of natural selection; Level 2 (L2): predicton "The fittest haplotype will be most frequent" but justification lacking a clear understanding of natural selection; Level 3 (L3) – prediction "The fittest haplotype will be most frequent" justified with important facts of natural selection (differential survival or differential reproduction); Level 4 (L4) - prediction "The fittest haplotype will be most frequent" justified with a clear understanding of the process of natural selection (differential survival and differential reproduction).

Figure 2 depicts the relative frequency of students' answers assigned to each level of evolution understanding in the entire sample (total sample average of evolution understanding=1.28). Without any specific instruction on evolution and natural selection, most of the students were classified at the L0 of evolution understanding. However, 17% of the students already reveal some understanding of natural selection (L3) and 6.4% of the students correctly predicted and justified their prediction describing the process of natural selection (L4). Kruskal-Wallis test reveals significant differences in the level of evolution understanding between classes ($p=0.001$), that were not explained by differences between classe grades ($p=0.241$, Mann Whitney U tests), but by differences between 4th grade classes ($p=0.001$,

Kruskal-Wallis tests). These results highlight the potential of this framework to detect statical differences in levels of evolution understanding between groups of students.

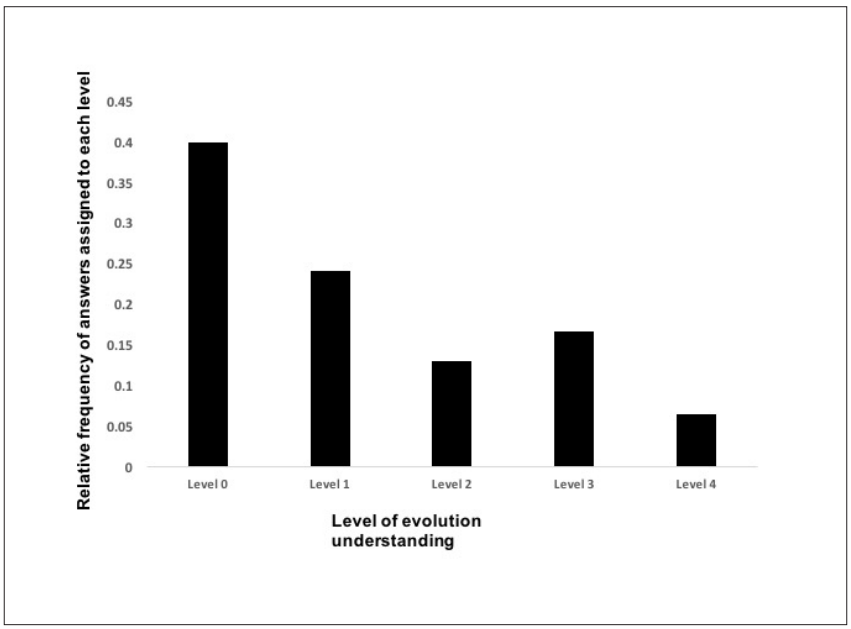

Fig. 2. Relative frequency of students' answers assigned to each level of evolution understanding in the entire sample.

CONCLUSIONS

We propose a framework (including tasks type, system of categories and level of understanding) to evaluate and characterize elementary students' understanding of evolution and natural selection, derived from patterns of answers observed in an empirical work with a sample of 3rd and 4th grade students. As observed in our sample, this framework allows to statistically identify differences in the level of understanding about evolution between classes, information that can be used to study the impact of developmental stages in evolution understanding and to design educational activities tailored to address students' needs in a specific class. It can also be used to test the impact of the educational activities, if pre and post test are used. Further studies are required to uncover potential problems and test additional opportunities in the use of this framework.

ACKNOWLEDGEMENTS

This work is financially supported by National Funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the project UID/CED/00194/2013. Xana Sá-Pinto is supported by Programa Operacional Capital Humano, Portugal 2020, European Social Fund and National Funds FCT/MEC (PIDDAC), through the SFRH/BPD/103613/2014 research grant.

REFERENCES

Berti, A.E., Barbetta, V., & Toneatti, L. (2017). Third-Graders' Conceptions about the origin of species before and after instruction: an exploratory study. *International Journal of Science and Mathematics Education*, *15*, 215-232.

Campos, R., & Sá-Pinto, A. (2013). Early evolution of evolutionary thinking: teaching evolution in elementary schools. *Evolution: education and outreach*, *6*, *25*.

Jenkins, K. (2009). Evolution in biology education: sparking imaginations and supporting learning. *Evolution: Education and Outreach*, *2*, 347-348.

Kelemen, D., Emmons, N.A., Schillaci, R.S., & Ganea, P.A. (2014). Young Children can be taught basic natural selection using a picture-storybook intervention. *Psychological Science, 25*(4), 893-902.

Nadelson, L., Culp, R., Bunn, S., Burkhart, R., Shetlar, R., Nixon, K., & Waldron, J. (2009). Teaching evolution concepts to early elementary school students. *Evolution: Education and Outreach*, 4, 267-274.

National Research Council (2012). A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas. Washington, DC: *The National Academies Press*.

Ministério da Educação (2004). Organização curricular e programas – 1º Ciclo. Mem Martins: *Departamento da Educação Básica.*

Wagler, R. (2010). A missing link: K-4 biological evolution content standards. *Evolution: Education and Outreach, 3*, 443-450.

McHugh, M.L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22* (3), 276-282.