

Part 14 / Strand 14

Evaluation and Assessment In Teaching and Learning

Co-editors: Jesper Bruun, Regula Grob & Olia Tsividanidou

Part 14 / Strand 14 Evaluation And Assessment in Teaching and Learning

Development, validation and use of standardised tests; instruments for measuring attitudes, interests, beliefs, self-efficacy, science process skills and conceptual understandings; authentic assessment, formative assessment, summative assessment, as well as approaches to assessment, monitoring student learning and implications for teaching. Includes international comparison studies such as TIMSS and PISA.

Sub-themes:

- 1) Development and Validation of Science Assessment Instruments
- 2) Formative Assessment Strategies in Science Education
- 3) Summative Assessment Practices in Science Classrooms
- 4) Implications of International Assessments (e.g., TIMSS, PISA) on Science Education
- 5) Assessing Student Attitudes and Beliefs in Science

Contents

Strand 14: Evaluation And Assessment In Teaching And Learning	1275
Assessment of Analogical Activities for Learning an Immune System Model	1276
Enhancing Climate Literacy Through Assessment Activities Designed by Pre-Service Primary Teachers.....	1285

Strand 14: Evaluation And Assessment In Teaching And Learning

Jesper Bruun¹, Regula Grob² and Olia Tsvidanidou³

¹University of Copenhagen, Denmark

²University of Teacher Education Lucerne, Switzerland

³Inquirium, Greece

Introduction

Evaluation and assessment are integral in contemporary Science Education. Together, they provide evidence that informs educational decision-making at multiple levels, from monitoring student learning to improving classroom practice and educational programmes. Within this broader landscape, assessment increasingly serves as a means of supporting learning, informing teaching, and facilitating the development of complex scientific competencies. As Science Education expands its focus towards modelling, systems thinking, sustainability, and other higher-order capabilities, assessment practices must also evolve to capture these multidimensional forms of learning. Increasing attention has been given to assessment approaches that make students' thinking visible, support learning throughout the educational process, and provide evidence of students' conceptual understanding and broader scientific competencies. The two contributions included in this strand illustrate this shift particularly well. Although they address different educational contexts and scientific domains, both investigate assessment as an integral component of learning. They demonstrate how carefully designed assessment tasks can reveal learners' reasoning, inform instructional decisions, and support the development of sophisticated scientific competencies.

Two themes connect the two contributions. The first theme is that assessing a complex competence begins with decomposition. Gómiz-Aragón and colleagues make secondary students' analogical reasoning assessable by resolving it into five operations, each rated on a four-level scale. Botzaki and Stavrou analyse how pre-service primary teachers design formative assessment by coding both the competency objectives the teachers target and the formative-assessment elements they deploy. In each study, the category system is the instrument through which competence is made visible.

The second theme has to do with the difficulty of reflection. In both studies the generative, evidence-gathering steps are handled competently, while the reflective moves are handled less surely. The students readily characterise the target and map similarities, but falter when asked to weigh the analogy's limitations and utility. The pre-service teachers elicit and even interpret evidence in varied ways, yet plan feedback, and opportunities for students to act on it, far less often, leaving the formative cycle unclosed.

Read together, the two studies span both sides of assessment—the student being assessed and the teacher learning to assess. We hope the ideas presented here prompt continued attention to how assessment can guide learning in a time when science education is increasingly concerned with sustainability and complex competencies.

Assessment of Analogical Activities for Learning an Immune System Model

Marta Gómiz-Aragón, María del Mar Aragón-Méndez, José María Oliva
University of Cádiz, Spain

This study investigates the use of analogies in secondary-school immunology teaching, comparing the immune system to a security system within a modelling framework. A total of 181 students participated in activities focused on five analogical operations: target characterisation, mapping of similarities, use of analogical language, identification of limitations, and recognition of utility. Regarding the first research question (RQ1), high levels of performance were observed in target characterisation and the mapping of similarities, whereas students encountered difficulties in using analogical language and identifying limitations. The perceived utility of the analogies varied considerably. For the second research question (RQ2), correlation and principal component analyses indicated that all five indicators contributed to a common construct of analogical performance. Nevertheless, the identification of limitations and the perceived utility showed weaker relationships with this construct. These findings suggest that constructing and reflecting upon analogies fosters both understanding of the immunity model and the development of skills in constructing and evaluating scientific models.

Keywords: Analogies in Science; Secondary Education; Teaching Processes

Introduction

Teaching immunology is essential for educating citizens capable of understanding information related to this field and making informed decisions that enhance their quality of life (Mixer et al., 2023). However, learning how the immune system functions presents significant challenges due to the large number of concepts involved and the multiple scales that must be considered, ranging from the molecular to the systemic and even the social level (Haidaris & Frelinger, 2019). Consequently, it is necessary to implement teaching strategies that facilitate the comprehension of these complex topics.

A particularly effective approach to addressing this challenge is through the use of analogies. Analogies act as mediators between the observable world and scientific models (Marzabal et al., 2024), allowing familiar phenomena to be related to complex scientific concepts. In this way, they support the construction of mental models (Clement, 2000) and assist in the creation, revision, and validation of scientific models (Gilbert & Justi, 2016). Within this context, analogies constitute a valuable resource for modelling-based approaches, as they enhance the understanding of school-level scientific models and promote engagement with the modelling process itself (Oliva & Aragón-Méndez, 2009).

This work forms part of a broader study analysing how students construct models of immunity using an analogy based on a security system. Specifically, the aim of the present study is to assess students' performance in analogical activities, thereby contributing to a better understanding of the role of analogies in modelling and how they can facilitate the comprehension of complex concepts in immunology.

Theoretical Framework

In science education, the use of models as learning objects and modelling as a scientific practice is well established. School-based models arise from the didactic transposition of scientific models, being simplified, restructured, and adapted to suit students' characteristics and the educational context (Coll et al., 2005; Treagust et al., 2002). Modelling, understood as the construction of models in the classroom, provides opportunities for students to engage in school-based scientific practices and plays an important role in their scientific development.

Analogies constitute a particular type of model, as each involves an underlying abstract representation that gives rise to the *analogy model*. This second-order model synthesises the relational structure between the *target* (what is to be explained) and the *analogue* (the known). Analogies are fundamental tools for teaching complex concepts, as they facilitate the understanding of unfamiliar ideas by relating them to familiar concepts that share similar structure or function (Gilbert & Justi, 2016; Tise et al., 2023). Traditionally, analogies have been used in science teaching as illustrative resources to explain isolated concepts. However, their more effective use requires articulating multiple ideas and relationships, thereby fostering the construction of structured and interconnected knowledge and avoiding fragmented or superficial learning.

Analogical reasoning represents a system of relational thinking based on identifying structural similarities and is considered one of the cognitive foundations of higher-order thinking (Eriksson et al., 2024; Richland & Simms, 2015). Its development promotes knowledge transfer and the construction of conceptual models (Coll et al., 2005; Gray & Holyoak, 2021). Through this process, students generate abstract conceptions from similarities between phenomena, supporting the understanding, application, and reconstruction of scientific models (Gilbert & Justi, 2016). Furthermore, analogies facilitate the practice and acquisition of scientific language, assisting students in articulating and explaining abstract concepts (Eriksson et al., 2024).

The approach based on identifying structural patterns across different domains is enhanced through modelling practices, which enable the generation of explanations, predictions, and inferences about the phenomena under study (Mozzer & Justi, 2018). In this process, students must assess the correspondences between the analogy and the scientific model, evaluating its validity and utility, and recognising when an analogy is appropriate and how to use it effectively (Petchev et al., 2023). Analogical thinking thus stimulates the construction, use, and evaluation of scientific models, while also fostering the acquisition of skills and epistemological perspectives essential to science and modelling (Oliva & Aragón-Méndez, 2017). The relationship between analogical thinking and modelling is bidirectional: constructing analogies involves, to some extent, building models in a domain familiar to students, while engaging in modelling processes enriches the use and meaning of analogies, promoting verbalisation, argumentation, and metacognitive reflection (Aragón-Méndez et al., 2014; Mozzer et al., 2011).

Thus, analogical reasoning, enhanced by modelling practices, not only contributes to the understanding of complex scientific concepts but also fosters advanced cognitive skills and competencies associated with knowledge construction, critical thinking, and scientific creativity (Zorlu & Zorlu, 2022).

For analogies to be effective in the classroom, careful planning is required, taking into account students' prior knowledge and potential risks of misunderstanding (Petchev et al., 2023). Teaching models such as *Teaching With Analogies* (TWA) by Glynn (1991) provide a structured framework to guide learning through analogies. The TWA model comprises six stages: (1) presentation of the target concept; (2) introduction of the analogue and activation of students'

creativity. (3) identification of relevant features in the analogue that show similarity with the target. (4) mapping of similarities and analysis of the compared systems. (5) drawing conclusions, generalising, and formulating hypotheses. (6) establishing the limits of the analogy and critically evaluating its relevance. This approach ensures effective knowledge transfer and promotes active student engagement, as students construct and refine their own analogies rather than receiving information passively (Saleem & Akbar, 2024).

Research in science education has explored, for decades, the cognitive processes involved in analogical reasoning, strategies for teaching with analogies, and factors affecting their effectiveness (Marrero & González, 2023). Analysing analogies created by students can reveal their conceptual understanding and reasoning processes (Tise et al., 2023). Given the multidimensionality of this process, student performance must be assessed using multiple indicators.

Nevertheless, questions remain regarding how to assess students' performance in analogy construction. Therefore, the present study addresses the following research questions:

RQ1: How does students' performance in constructing the immune system analogy vary across different performance indicators?

RQ2: To what extent do these indicators converge and form part of a single construct?

Research Design

Didactic Context

The study involved 181 students aged 14–15 (3rd year of secondary education) from two public schools in Spain. It was embedded in a teaching–learning sequence based on modelling. In this sequence, a security system analogy was employed to facilitate the learning of the immunity and vaccination model at the microscopic level, with a focus on the cellular processes of the immune system.

In immunology teaching, the defence or warfare system analogy is commonly used to represent the immune system (cf. Niebert et al., 2012). The security system is an appropriate analogue due to functional similarities: both detect threats, respond to them, and store information to prevent future attacks. To reinforce the analogy, a silent animation adapted from Bryce (2018) was used. In this animation, pathogens were represented as criminals and immune cells as security agents. Databases and emergency services were also included to strengthen functional correspondences. The absence of audio allowed students to interpret the animation with teacher guidance, promoting interaction, dialogue, and active model construction.

Two activities related to the analogy were analysed: its construction and metacognitive reflection on the analogical model.

In the first activity, students constructed the analogy in two steps. First, they mapped the similarities between the target (immune system) and the analogue (security system). Second, they drew conclusions regarding the characteristics of the target. The analysis focused on the language used, distinguishing analogies from metaphors. Analogies establish a systematic correspondence between the two domains, whereas metaphors transfer meaning from the source to the target (Gilbert & Justi, 2016).

The analogy was interpreted based on Gentner's (1982) structural mapping theory. According to this theory, students establish systematic correspondences between objects in the analogue domain (criminals, agents, databases) and the target domain (pathogens, leukocytes, immunological memory). These correspondences are based on the functional relationships

connecting the two domains. This approach preserves coherent systems of relations and prioritises higher-order structural connections, facilitating the construction of mental models of the immune system and enabling understanding of complex processes beyond the superficial attributes of the elements. Students worked individually with an incomplete analogy. They were required to identify correspondences and relationships between the security system and the immune system.

The second activity focused on fostering students' metacognition regarding the analogical model. Students reflected on the model's usefulness for understanding immune system mechanisms and its limitations. This promoted critical evaluation of the analogical model itself. Guiding questions included: *To what extent does the analogical model allow understanding of how the immune system functions? What are the limitations of the model? To what extent can immune cells be compared to human agents in a security system?*

The aim was for students to understand that scientific models are approximate representations of reality. Models have limitations, and a single phenomenon can be explained using different models. Furthermore, participating in the reconstruction and evaluation of the analogical model allows students to gain a deeper understanding of their own learning process. It also supports the development of competencies related to modelling, analogical thinking, and metacognitive reflection (Gray & Holyoak, 2021).

Overall, students were expected to engage in five analogical operations. Three of these correspond to the TWA model: identifying relevant features of the target, mapping similarities, and indicating the limitations of the analogy. Two additional operations were included: the use of analogical language and metacognitive processes related to the analogy's usefulness in learning the immunity and vaccination model.

Analysis Method

With respect to the five analogical operations identified in the activities—characterisation of the attributes of the target domain, mapping of similarities between the analogue and the target, use of analogical language, identification of the analogy's limitations, and recognition of the model's usefulness for learning—four performance categories were established. These categories allowed the assessment of students' levels of understanding and application (Table 1). The categorisation system was agreed upon through inter-rater consensus to ensure consistency in the classification of responses. Inter-rater reliability was verified using Cohen's kappa coefficient, appropriate for ordinal categorical variables, thereby ensuring consistent classification of students' responses.

Table 1. Category system for performance in analogical activities.

Indicator	Category 1	Category 2	Category 3	Category 4
Target (T). Characterization of the target domain attributes	No correspondence with model elements.	Names immune elements or mechanisms but without description.	Describes immune elements or mechanisms but without naming them.	Fully characterizes immune elements or mechanisms.
Similarities (S). Mapping similarities	No similarities established.	Inadequately established.	Correct but incomplete mapping.	Fully accurate mapping.
Language (A). Use of analogical language	No comparative structures used.	Simple comparative structures.	Metaphorical language without full analogical comparison.	Fully adequate comparative structures.
Limitations (L). Identification of analogy limitations	No limitations identified.	Superficial or unspecific limitations.	Structural limitations identified.	Functional limitations identified.
Utility (U). Recognition of the utility of the analogy for learning the model	No assessment of utility.	Negative assessment.	Both positive and negative assessment.	Positive assessment.

To address RQ1, a descriptive analysis was conducted. The frequency distribution for each category was calculated for each operation, with the aim of identifying patterns and trends in students' responses.

Subsequently, to address RQ2, the relationships between the different indicators were analysed. Kendall's tau-b coefficient, suitable for ordinal variables, was employed. Additionally, a categorical principal component analysis was carried out, allowing exploration of whether the different indicators converge into a single construct of analogical performance.

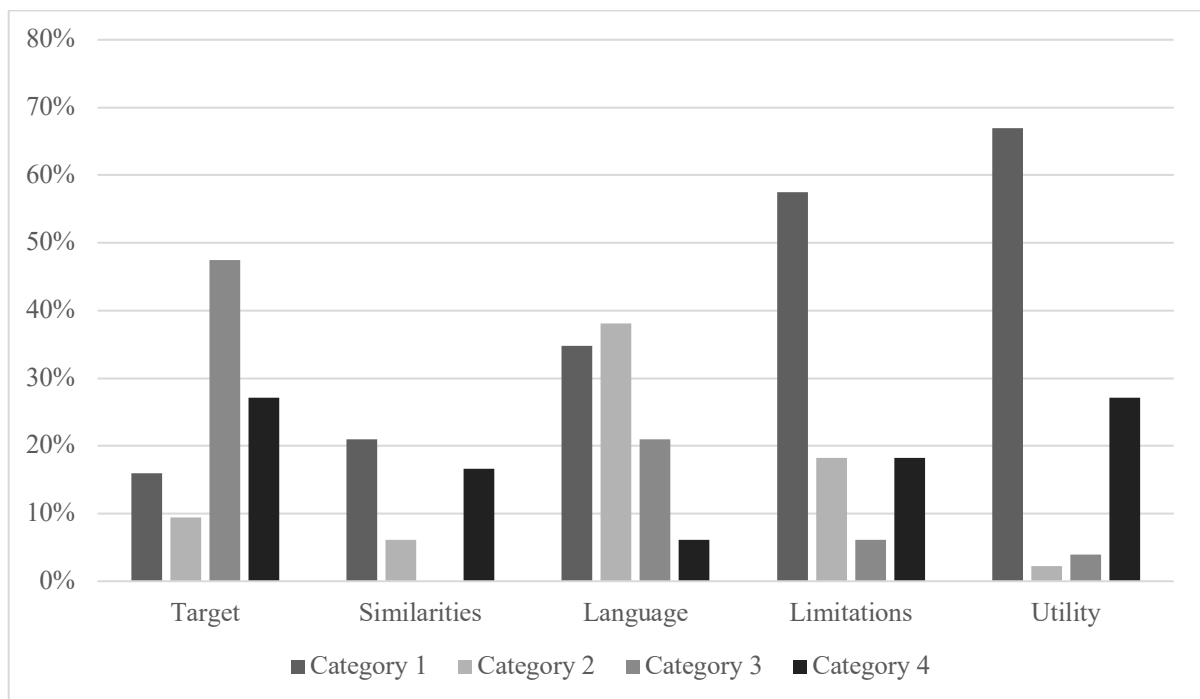
All analyses were performed using IBM SPSS Statistics v.29. This procedure ensured the validity and reliability of the results and provided a robust framework for interpreting students' performance in the analogical activities.

Results

Assessment Of Performance In The Construction Of The Analogy

In response to RQ1, which examined students' performance in constructing the immune system analogy, Figure 1 presents the frequency distribution of performance levels across the five analogical activities.

Figure 1. Patterns in performance on analogical activities



The highest levels of performance were observed in the mapping of similarities and in the use of analogical language, whereas identification of limitations emerged as the weakest area, with many students failing to reflect on the boundaries of the analogical model.

By contrast, recognition of the model's usefulness exhibited marked polarisation: 27.1 % of students positively valued the effectiveness of the analogy for understanding the model, while the majority, with few exceptions, did not perceive its applicability. This pattern indicates a limited degree of reflection on the model's usefulness.

Table 2 presents representative examples of students' responses across the different categories for the five indicators of the analogical operations. Reference responses correspond to the expected answers for each indicator, while the subsequent columns display actual student responses, selected to illustrate performance patterns and the difficulties observed.

Table 2. Representative Examples of Student Responses

Indicator	Categorised examples from student portfolios
Target	Reference: Immunological memory allows a faster and more effective secondary response upon a new infection by the same pathogen. Category T1: Antibodies detect the microorganism and reject it Category T2: Having fought that pathogen, they are prepared for another attack Category T3: Lymphocytes remember the pathogen for when it returns and know how to eliminate it quickly. Category T4: Immunological memory allows lymphocytes to produce specific antibodies, so that if a second encounter occurs, they neutralise it faster.
Similarities	Reference: In both systems, there are elements (phagocytes and riot police) capable of stopping threats (pathogens and criminals). Category S1: The systems act like a threat. Category S2: Both have transport. Category S3: They try to eliminate the threat. Category S4: In both cases, responses are general and rapid.
Language	Reference: Just as different security agents protect populations from various criminals, white blood cells are responsible for protecting the organism from infections. Category A1: For identifying. Category A2: Both defend the place. Category A3: Both systems have elements that watch and inspect for threats. Category A4: Both systems have mechanisms that help detect and eliminate threats.
Limitations	Category L1: It explains very well how the immune system works. Category L2: The immune system can be compared with other systems; although similar, each has its own characteristics. Category L3: The police model and the immune system are similar but differ in important aspects: a single microorganism can cause an infection, whereas a single criminal does not represent a great threat; microorganisms act independently, without a leader. Category L4: The model is very useful but cannot literally represent reality: we will not see phagocytes fighting with weapons. The comparison has limits between the biological and the social or ethical domains.
Utility	Category U1: Knowing the terms and having learned a bit more about the topic, I can understand the model and why many things work as they do. Without prior knowledge, I would not have understood anything, as the terms are not fully explained. Category U2: The model does not explain the order in which different types of immune cells act. Category U3: Models are essential to understand reality, but they must be updated with new discoveries, as they do not reproduce it completely. Category U4: The analogy facilitates understanding by relating the immune system to something familiar, although its abstract and non-visible nature may hinder comprehension.

Note. The students' original responses were in Spanish. The examples presented in this table have been translated into English for clarity and consistency.

Convergence Of Indicators

In response to RQ2, Table 3 presents the correlation matrix calculated using Kendall's tau-b coefficient. The Target and Similarities indicators exhibit the strongest relationship among all pairs of variables. The Language indicator shows significant correlations with all other variables, although these are of smaller magnitude compared with Target and Similarities. The Limitations and Utility indicators display weaker correlations, yet they remain statistically significant. This suggests that these dimensions are less directly related to performance in constructing the analogy.

Table 3. Kendall's tau-b correlation matrix.

	Target	Similarities	Language	Limitations	Utility
Target	1,000				
Similarities	0,616***	1,000			
Language	0,434***	0,569***	1,000		
Limitations	0,336***	0,339***	0,139*	1,000	
Utility	0,359***	0,303***	0,183**	0,229***	1,000

* $p < .05$; ** $p < .01$; *** $p < .001$

Finally, a categorical principal component analysis for ordinal variables was performed (Table 4). This analysis revealed a single-factor solution explaining 54.1 % of the variance. This indicates that all five indicators contribute to a common construct, despite differing levels of difficulty and assessment methods. The extracted factor is particularly associated with the characterisation of the target domain and the mapping of similarities. In contrast, the identification of limitations and the evaluation of utility are more weakly connected to the other indicators.

Table 4. Factor loadings from nonlinear principal component analysis.

	Factor 1
Target	0,867
Similarities	0,886
Language	0,736
Limitations	0,583
Utility	0,538

Discussion, Conclusions, and Implications

Although the literature provides extensive exploration of the use of analogies as a teaching resource in science education, few studies have assessed students' performance in analogy-based activities. This study seeks to contribute to this area. The results indicate that the analogical operations involved in an activity comparing the immune system to a security system, designed according to the TWA model, allow for a comprehensive evaluation of students' performance in completing these tasks.

Students achieved medium to high performance in characterising the attributes of the immune system and in mapping similarities with a security system. However, their performance in using analogical language was lower, possibly due to a general tendency towards laxity in written expression (Archila, 2015).

Difficulties were also observed in analysing the limitations of the analogy, suggesting that this task is more complex than identifying similarities. This complexity relates to the potential of analogies to promote critical thinking, as learning occurs not only by establishing correspondences but also by questioning them and recognising their boundaries (Heywood, 2002; Haglund, 2013). Therefore, providing scaffolding to support the identification of analogy limitations could further enhance students' critical capacity. In this process, analogical reasoning facilitates higher-order thinking, such as the evaluation of ideas and the generation of inferences (Richland & Simms, 2015; Zorlu & Zorlu, 2022).

Although students often noted the analogy's usefulness for learning about immunity, this perception was not clearly reflected in the assessment instruments. When asked about the analogy's usefulness for understanding the immune system model, many students merely used the analogy to demonstrate conceptual understanding, without explicitly indicating whether it had been helpful. This ambiguity suggests a limitation in the formulation of the question.

For future implementations, rather than posing the open question, "To what extent does the analogical model allow you to understand how the immune system works?", it may be useful to reword the question and provide prompts to guide reasoning, such as: "Does the analogy allow you to explain what happens when a pathogen enters the body?" or "Can it be used to predict what would happen if another type of substance were introduced?"

Statistical analysis suggests an underlying factor accounting for students' performance in reconstructing the immune system–security system analogy. This factor may reflect the impact of the activities on developing analogical thinking, linking analogical use with modelling

competencies. Both share cognitive operations –such as identifying structural correspondences, drawing inferences, and evaluating representations– that support model construction and revision (Aragón-Méndez & Oliva, 2020; Justi & Gilbert, 2002). This aligns with Lin and Chao (2024), who define analogical modelling competence as the ability to use analogical reasoning to construct models, integrating cognitive, metacognitive, and behavioural aspects. Classroom use of analogies fosters this competence by offering opportunities to construct and evaluate models.

In line with this, the findings of the present study suggest that students' performance in reconstructing the immune system analogy may reflect not only progress in understanding the immunity and vaccination model but also advancement in their competence to construct and revise scientific models. This hypothesis will be explored further in future research.

Acknowledgements

This study was supported by Grant PID2022-136353NB-I00, funded by MICIU/AEI/10.13039/501100011033 and co-financed by the ERDF/EU.

References

- Aragón-Méndez, M. M., & Oliva, J. M. (2020). Analogías, simulaciones y experimentos mentales para la construcción del modelo del cambio químico [Analogies, simulations and thought experiments in order to construct the chemical change model]. *Educació Química*, 27, 35–41. <https://doi.org/10.2436/20.2003.02.204>
- Aragón-Méndez, M. M., Oliva, J. M., & Navarrete, A. (2014). Contributions of learning through analogies to the construction of secondary education pupils' verbal discourse about chemical change. *International Journal of Science Education*, 36(12), 1960–1984. <https://doi.org/10.1080/09500693.2014.887237>
- Archila, P. A. (2015). Uso de conectores y vocabulario espontáneo en la argumentación escrita: aportes a la alfabetización científica [Use of connectors and spontaneous vocabulary in written argumentation: contributions to scientific literacy]. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 12(3), 402–418. https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2015.v12.i3.02
- Bryce, E. (2018). *How does the immune system work?* <https://ed.ted.com/lessons/how-does-the-immune-system-work-emma-bryce>
- Clement, J. (2000). Model based learning as a key research area for science education. *International Journal of Science Education*, 22(9), 1041–1053. <https://doi.org/10.1080/095006900416901>
- Coll, R. K., France, B., & Taylor, I. (2005). The role of models/and analogies in science education: implications from research. *International Journal of Science Education*, 27(2), 183–198. <https://doi.org/10.1080/0950069042000276712>
- Eriksson, S., Gericke, N., & Thörne, K. (2024): Analogy competence for science teachers, *Studies in Science Education*, 1-29. <https://doi.org/10.1080/03057267.2024.2434797>
- Gilbert, J. K., & Justi, R. (2016). Analogies in modelling-based teaching and learning. In J. K. Gilbert & R. Justi (Eds.), *Modelling-based teaching in science education* (pp. 149–169). Springer. https://doi.org/10.1007/978-3-319-29039-3_8
- Glynn, S. M. (1991). Explaining science concepts: A teaching with analogies model. In S. M. Glynn, R. H. Yearny, & B. K. Britton (Eds.), *The psychology of learning science* (pp. 219–240). Lawrence Erlbaum.
- Gray, M., & Holyoak, K. J. (2021). Teaching by analogy: From theory to practice. *Mind, Brain, and Education*, 15(2). <https://doi.org/10.1111/mbe.12288>
- Haglund, J. (2013). Collaborative and self-generated analogies in science education. *Studies in Science Education*, 49(1), 35–68. <https://doi.org/10.1080/03057267.2013.801119>
- Haidaris, C. G., & Frelinger, J. G. (2019). Inoculating a new generation: Immunology in medical education. *Frontiers in Immunology*, 10, 2548. <https://doi.org/10.3389/fimmu.2019.02548>
- Heywood, D. (2002) The Place of Analogies in Science Education, *Cambridge Journal of Education*, 32(2), 233–247. <https://doi.org/10.1080/03057640220147577>
- Justi, R. S., & Gilbert, J. K. (2002). Modelling, teachers' views on the nature of modelling, and implications for the education of modellers. *International Journal of Science Education*, 24(4), 369–387. <https://doi.org/10.1080/09500690110110142>

- Lin, J. W., & Chao, H. Y. (2023). Developing an instrument to examine students' analogical modeling competence: An example of electricity. *Science Education*. <https://doi.org/10.1002/sce.21828>
- Marrero, J. J., & González, P. (2023). Investigaciones sobre el uso de analogías en el aula de ciencias: una revisión sistemática [Investigations about the use of analogies in science classrooms: A systematic review]. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 20(1). https://doi.org/10.25267/Rev_Eureka_ensen_divulg_cienc.2023.v20.i1.1101
- Marzabal, A., Merino, C., Soto, M., & Cortés, A. (2024). Modeling-based science education. In A. Marzabal y C. Merino (eds.), *Rethinking science education in Latin-America: Diversity and equity for Latin American students in science education* (pp. 245-261). Springer.
- Mixter, P. F., Kleinschmit, A. J., Lal, A., Vanniasinkam, T., Condry, D. L. J., Taylor, R. T., Justement, L. B., & Pandey, S. (2023). Immune Literacy: a Call to Action for a System-Level Change. *Journal of Microbiology & Biology Education*, 24(1). <https://doi.org/10.1128/jmbe.00203-22>
- Mozzer, N. B., & Justi, R. (2018). Modelagem analógica no ensino de ciências [Analogical modeling in science education]. *Investigações em Ensino de Ciências*, 23(1), 155–182. <https://doi.org/10.22600/1518-8795.ienci2018v23n1p155>
- Mozzer, N., Justi, R., & Costa, P. (2011). *Students' analogical reasoning when participating in modelling-based teaching activities*. 9th International Conference of the European Science Education Research Association, Lyon, France.
- Niebert, K., Marsch, S., & Treagust, D. F. (2012). Understanding needs embodiment: A theory-guided reanalysis of the role of metaphors and analogies in understanding science. *Science Education*, 96(5), 849-877. <https://doi.org/10.1002/sce.21026>
- Oliva, J. M., & Aragón-Méndez, M. M. (2009). Contribución del aprendizaje con analogías al pensamiento modelizador de los alumnos en ciencias: marco teórico [Contribution of learning with analogies to students' modelling thinking in science: Theoretical framework]. *Enseñanza de las Ciencias*, 27(2), 195–208. <https://doi.org/10.5565/rev/ensciencias.3731>
- Oliva, J. M., & Aragón-Méndez, M. M. (2017). Modelización y pensamiento analógico en el aprendizaje del cambio químico [Modeling and analogical thinking in learning chemical change]. *Revista Brasileira de Pesquisa em Educação em Ciências*. <https://doi.org/10.28976/1984-2686rbpec2017173903>
- Petchey, S., Treagust, D., & Niebert, K. (2023). Improving university life science instruction with analogies: Insights from a course for graduate teaching assistants. *CBE—Life Sciences Education*, 22(2). <https://doi.org/10.1187/cbe.22-07-0142>
- Richland, L. E., y Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(2), 177-192. <https://doi.org/10.1002/wcs.1336>
- Saleem, A., & Akbar, R. A. (2024). Role of Analogy Based Teaching to Enhance Higher Order Skills in Chemistry Learning at Grade IX. *Human Nature Journal of Social Sciences*, 4(2), 372-388.
- Tise, J. C., Sperling, R. A., Dann, M. S., & Young, T. M. (2023). Teaching postsecondary students to use analogies as a cognitive learning strategy: An intervention. *CBE—Life Sciences Education*, 22(1). <https://doi.org/10.1187/cbe.22-05-0084>
- Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education*, 24(4), 357-368. <https://doi.org/10.1080/09500690110066485>
- Zorlu, F., & Zorlu, Y. (2022). The investigation of preservice science teachers' analogical reasoning skills from thinking skills. *Thinking Skills and Creativity*, 43, 100981. <https://doi.org/10.1016/j.tsc.2021.100981>

Enhancing Climate Literacy Through Assessment Activities Designed by Pre-Service Primary Teachers

Eleni Botzaki and Dimitris Stavrou
University of Crete, Greece

Addressing Climate Change (CC) challenges requires fostering climate-literate students by developing CC competencies, which can be strengthened through integrated assessment in CC teaching. However, a significant lack of guidance on assessing CC competencies highlights the need for meaningful support to teachers in effectively implementing assessment strategies. The present study examines how nineteen pre-service primary teachers design and develop activities to assess CC competencies. Conducted in two phases, the study first introduced participants to key CC concepts, aspects of climate literacy, and assessment principles. Subsequently, they designed and developed CC teaching scenarios, with a particular emphasis on integrating assessment dimension to foster climate literacy among students.

The findings shed light on the pre-service primary teachers' choices regarding the CC competencies that they aim to foster through their scenarios and the structure of formative assessment elements they employ for assessing these competencies. In more detail, they place particular emphasis on assessing Systems Thinking. Additionally, although participants develop their activities to elicit evidence of students' progress, their feedback practices remain limited, indicating a gap in their formative assessment design. Thus, these findings stress the importance of enhancing assessment practices in teacher education to foster climate literacy.

Keywords: Formative Assessment, Pre-Service Teacher Education, Sustainability

Introduction

Climate Change (CC) is one of the most crucial environmental challenges, mainly caused by anthropogenic activities that harm the environment, affecting multiple aspects of human life (IPCC, 2023). Addressing CC issues requires the preparation of citizens, capable of mitigating the impacts of CC and partially adapting to them. CC education needs to foster climate literacy in students by developing CC competencies - a combination of knowledge, skills, and attitudes (Bianchi et al., 2022). Various frameworks were developed to define the necessary competencies, with a general focus on sustainability, as climate change is inherently part of the broader domain of sustainability. Widely recognized frameworks include Wiek et al. (2011) and the European GreenComp framework (Bianchi et al., 2022), identifying competencies such as systems thinking, futures thinking, and action-taking. Nevertheless, to address key aspects of climate literacy more effectively, recent efforts focus on defining CC competencies using broader sustainability frameworks, such as that of Taurinen et al. (2024).

In this context, achieving the complex goal of developing CC competencies requires contemporary teaching approaches emphasizing active learning (UNESCO, 2017). Such approaches view teaching and assessment as interconnected processes that together enhance student learning (Harris et al., 2019). Accordingly, focusing solely on the final outcome and knowledge acquisition, as is common in traditional assessment methods (e.g., summative assessment), is insufficient to capture the complexity of CC competencies (Annelin & Boström, 2022; Redman et al., 2021). Thus, the assessment of CC competencies requires methods that focus on the whole learning process and provide feedback (Grangeat et al., 2021), such as formative assessment (Annelin & Boström, 2022). The implementation of formative assessment in classroom contexts is based on five key elements: a. setting and sharing with students clear learning objectives and assessment criteria; b. eliciting information from students; c. interpreting

the elicited information; d. providing constructive feedback; and e. creating opportunities for students to apply feedback. These five elements form a cyclical process that supports ongoing learning (e.g., Dini et al., 2020; Grob et al., 2019; Black & Wiliam, 2018; Haug & Ødegaard, 2015).

Nevertheless, teachers face challenges when integrating assessment into their teaching, especially when dealing with multidimensional learning goals, such as the development of competencies (Grangeat et al., 2021). These challenges vary across the steps of the formative assessment cycle, as each step involves different teacher decisions and practices. First of all, teachers tend to focus mainly on assessing knowledge-related objectives rather than students' skills or attitudes (Ropohl & Rönnebeck, 2019; Cisterna & Gotwals, 2018;). In addition, although the literature highlights the importance of co-designing learning goals with students, teachers often avoid setting learning goals collaboratively or explicitly sharing them with students (Cisterna & Gotwals, 2018). Another challenge concerns the elicitation of evidence, in order to assess students' performance and progress in relation to the learning objectives. In particular, when the learning objectives refer to competency development, assessing the conceptual knowledge is not sufficient. Instead, teachers need evidence of students' thinking and reasoning, as well as their skills and attitudes (Grangeat et al., 2021). Thus, a wide variety of evidence is required in order to build a more holistic picture of students' progress (Grob et al., 2021; Cisterna & Gotwals, 2018). Besides the elicitation of evidence, teachers - especially pre-service teachers - often face difficulties during the interpretation of the collected evidence. Particularly, they tend to interpret the elicited information more superficially (Ropohl & Rönnebeck, 2019), either because they do not consistently connect it to learning objectives or because they do not examine in depth how students understand, reason, and argue. This, in turn, limits their ability to identify students' actual learning needs (Ropohl & Rönnebeck, 2019; Ateh, 2015). Consequently, limited interpretation of evidence can prevent teachers' instructional decision-making, particularly about designing the appropriate next steps and providing feedback that moves learning forward (Ropohl & Rönnebeck, 2019). Regarding teachers' actions after the interpretation of evidences, research shows that their decisions often do not meaningfully support the teaching and learning process. Many teachers face challenges in providing effective feedback. When feedback is given, it is often general, limited in depth, and non-individualized (Ropohl & Rönnebeck, 2019). Teachers also tend to focus mainly on whether students' answers are correct, rather than responding to students' thinking in ways that support deeper learning (Ateh, 2015). As a result, formative assessment cycles often remain incomplete and, consequently, may not effectively promote students' learning (Cisterna & Gotwals, 2018).

Overall, both pre- and in-service teachers need further support in order to use formative assessment for the development of competencies. At the same time, research on assessing CC competencies remains limited, particularly with respect to formative assessment in primary education (e.g., Cebrian et al., 2025). There is also a lack of guidance for teachers on how to use assessment tools effectively, which highlights the need for meaningful support (Redman et al., 2021).

Based on the above, the study aims to explore how pre-service teachers design and develop activities for assessing CC competencies. More specifically, emphasis is given to the phase of the planning of the formative assessment, which is considered as a critical component of its effectiveness (e.g., Grangeat et al., 2021). Accordingly, the main research question of the study is, "*How do pre-service teachers design assessment sequences to support the development of the CC competencies?*", which is analyzed through two sub-questions:

- What CC competency objectives do pre-service teachers set to assess in their activities?

- How do pre-service teachers structure the elements of formative assessment?

Methods

Empirical Study Description

This study involves nineteen pre-service primary teachers, who participated in a 13-week seminar course of the Department of Primary Education at the University of Crete, developing teaching scenarios about the development of CC competencies with the integration of assessment dimension. The empirical study is divided into two phases, as presented in Table 1.

Table 1. Description of CC competencies categories.

Empirical Study Phases	Meetings	Content
Phase A <i>Participants' Preparation</i>	<i>1st</i>	<i>CC scientific concepts</i>
	<i>2nd - 3rd</i>	<i>CC competencies</i>
	<i>4th</i>	<ul style="list-style-type: none"> • <i>Assessment of CC competencies</i> • <i>Formative assessment strategies</i>
	<i>5th - 6th</i>	<ul style="list-style-type: none"> • <i>Framework for the design of assessment tools and activities</i> • <i>Examples of assessment activities and tools</i>
Phase B <i>Design & Development of Teaching Scenarios</i>	<i>7th</i>	<i>Setting Learning Objectives for Whole Teaching and Individual Activities</i>
		<i>Reflection & Personalized feedback (per group)</i>
	<i>8th - 9th</i>	<ul style="list-style-type: none"> • <i>Refinement of the Learning Objectives</i> • <i>Activities & Tools Design: Determining the indicators to monitor learner progress</i>
		<i>Reflection & Personalized feedback (per group)</i>
	<i>10th</i>	<i>Activities & Tools Design: Determining the activity characteristics (e.g., task format)</i>
		<i>Reflection & Personalized feedback (per group)</i>
	<i>11th</i>	<i>Development of activities and assessment tools (plus Worksheets)</i>
		<i>Reflection & Personalized feedback (per group)</i>
	<i>12th</i>	<i>Teaching scenario refinement</i>

During **Phase A** (six meetings), participants were introduced to a. core concepts of CC (e.g., drivers, causes) and climate literacy, b. principles for assessing CC competencies, c. formative assessment strategies, and d. specific examples of designing and utilizing assessment tools and activities to foster CC competencies. Then, in **Phase B**, pre-service teachers developed teaching scenarios, emphasizing the integration of the assessment tools and activities to foster CC competencies. Six plenary meetings were thus dedicated to supporting participants during the development process. Additionally, four individual meetings were held with each group between

the main meetings to address challenges, provide personalized feedback, and support reflection on their teaching scenario.

Data Collection

Data were collected through participants' developed teaching scenarios (n=7), including the assessment activities and tools that they designed. Due to the nature of the research, qualitative methods of content analysis were used (Bryman, 2016). Specifically, the category systems were developed based on the relevant literature and were refined or adjusted in response to the empirical data.

Data Analysis

Data analysis examines participants' teaching scenarios along two main axes: a. CC competencies objectives, and b. elements of formative assessment. Regarding the first sub-research question, a categorization scheme for the CC competency objectives was developed based on literature (e.g., Taurinen et al., 2024; Bianchi et al., 2020; Wiek et al., 2011) and empirical data, as presented in Table 2.

Table 2. Description of objectives for CC competencies.

Categories	Criteria
Systems Thinking	<ul style="list-style-type: none"> • Complex systems' features (e.g. feedback loops) • Various domains (e.g. society, economy) • Effects of actions/choices
Futures Thinking	<ul style="list-style-type: none"> • Time scale <ul style="list-style-type: none"> ◦ Real data (past-present) ◦ Trends (future) • Future scenarios development
Strategic Competency	<ul style="list-style-type: none"> • Fostering action-taking • Design of solutions/action plans
Normative Competency	<ul style="list-style-type: none"> • Evaluation of the sustainability of current or future states • Sustainability values • Decision-making (e.g., addressing dilemmas)
Interpersonal Competency	<ul style="list-style-type: none"> • Evaluation of diverse perspectives • Communication, negotiation, and collaboration skills

Additionally, each activity's objectives were examined to identify the specific dimensions of CC competencies—knowledge, skills, and attitudes. Specifically, **knowledge objectives** focus on understanding and elaborating principles, theories, and facts about scientific core concepts and phenomena of CC. **Skills objectives** emphasize applying knowledge and expertise to accomplish tasks, fostering cognitive and practical skill development. **Attitude objectives** involve forming personal values, attitudes, and intentions that influence individuals' behavior.

Regarding the second research sub-question, the assessment activities and tools included in the participants' teaching scenarios were analyzed to identify how formative assessment elements were structured and combined across the scenarios. For this purpose, a categorization scheme (see Table 3) was applied, which was based on formative assessment elements described in the

literature (e.g., Dini et al., 2020; Grob et al., 2019; Black & Wiliam, 2018) and refined through the empirical data.

Table 3. Description of Formative Assessment Elements.

Categories	Description
Using Assessment Criteria	Specific criteria are used to assess the students' performance
Eliciting Information	Activities designed to capture aspects of competencies – knowledge, skills, attitudes.
Interpreting Information	Interpreting the collected evidence to identify students' progress and needs in developing CC competencies.
Providing Feedback	Providing information/comments to support CC competencies development and enhance learning outcomes.
Applying Feedback	The instructional design supports the direct use of feedback in following tasks.

Findings

CC competencies objectives

To address the first research sub-question about the CC competency objectives that pre-service teachers set to assess in their activities, Tables 4-5 summarize the findings.

According to the analysis of the participants' teaching scenarios, Table 4 presents the absolute frequency of the CC competencies objectives that addressed in the assessment activities developed by the pre-service teachers. Most learning objectives focus on Systems Thinking (n=30), while the other competencies were covered less frequently, especially Interpersonal Competency. This indicates a priority given to assessment of Systems Thinking over the other competencies.

Table 4. Learning objectives for CC competencies.

Categories	Absolute Frequency
Systems Thinking	30
Futures Thinking	14
Strategic Competency	17
Normative Competency	17
Interpersonal Competency	10

Additionally, Table 5 presents the absolute frequency of the different dimensions of CC competencies, demonstrating the tendency of participants to primarily assess the cognitive dimension, while indicating an effort to include skills assessment (n=22). Assessing values seems challenging for participants, resulting in a considerably lower frequency of values-related objectives (n=12) compared to the other two dimensions.

To answer the second research sub-question on how pre-service primary teachers structure formative assessment elements, Tables 6-7 and Figure 1 were created to summarize the findings.

Table 5. Dimensions of CC Competency Objectives.

Categories	Frequency
Knowledge	27
Skills	22
Attitudes	12

Formative Assessment Elements

Specifically, Table 6 presents how many times participants designed each of formative assessment elements. Participants used only assessment criteria that they designed themselves without considering engaging students in this process. These criteria were typically presented as checklists or detailed rubrics. Regarding “eliciting information”, participants recognized its importance and designed different types of elicitation tasks beyond whole-class discussion or simple questioning, including tasks requiring the creation of artifacts (e.g., concept maps, leaflets, posters) and demonstration of skills (e.g., diagram construction, action-plan design). In addition, participants most of the times did not plan just to collect data, but also to make sense of it, identifying points in their scenarios where interpretation would occur, based either on predetermined criteria or on more intuitive judgement. They also specified if evidence would be interpreted by the teacher or by students through self-assessment and peer-assessment. However, although the element “interpreting information” was frequently mentioned, feedback based on these interpretations was planned to be given much less often, indicating a difficulty for pre-service primary teachers. They organised the provision of feedback through multimedia (e.g., video, images), rubrics and individual or plenary discussions. Finally, “applying feedback” was the least frequently designed element by participants, indicating that feedback rarely leads to follow-up actions. In the few cases that participants integrated this element in their teaching scenarios, they focused on creating opportunities for students to revise their answer, artifact or their performance on the task.

Table 6. Elements of Formative Assessment.

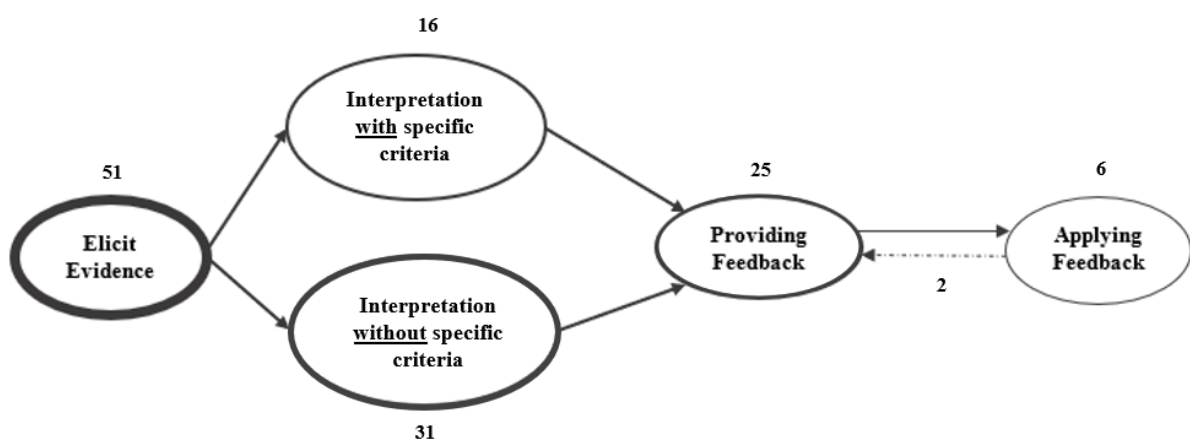
Categories	Frequency
Using Assessment Criteria	16
Eliciting Information	51
Interpreting Information	47
Providing Feedback	27
Applying Feedback	6

Table 7 provides a more detailed view by breaking down participants’ choices of formative assessment elements by CC competency. While the previous table summarized the elements used in general, Table 7 shows the pattern of element selection within each competency. Reading across the rows, the table indicates that complete formative assessment cycles were only attempted for Systems Thinking and Futures Thinking. The findings also reveal that pre-service teachers may consider Systems Thinking a demanding competency for both themselves and their students and they tried to better support its development. Accordingly, most formative assessment elements appear more frequently in Systems Thinking, particularly “Providing Feedback”.

Table 7. Formative Assessment Elements across CC Competencies.

FA Elements CC Competencies	Using Assessment Criteria	Eliciting Information	Interpreting Information	Providing Feedback	Applying Feedback
Systems Thinking	8	18	17	14	4
Futures Thinking	2	7	7	3	2
Strategic Competency	3	10	9	3	0
Normative Competency	1	10	8	5	0
Interpersonal Competency	2	6	6	2	0

Furthermore, the analysis of teaching scenarios revealed that participants structured the formative assessment elements that were examined earlier (see Table 6 & 7), forming integrated assessment pathways within activities. These pathways were illustrated for each CC competency, with Figure 1 presenting a comprehensive pathway summarizing all CC competencies, where line thickness reflects the frequency of each strategy and the dashed line indicates a return to a previous phase. However, providing feedback is challenging for pre-service teachers, who do not often incorporate it into their activities. When included, it is typically given at the end and does not provide students with the opportunity to apply it constructively.

Figure 1. Comprehensive assessment pathway of all CC competencies.

Discussion & Conclusions

The findings show how pre-service teachers planned formative assessment elements within teaching scenarios developed to support the development of CC competencies. Overall, they primarily targeted knowledge-related objectives, with a particular emphasis on Systems Thinking. The findings also reveal challenges in relation to feedback practices.

More specifically, pre-service teachers designed activities aimed at fostering climate literacy, with a particular emphasis on assessing the cognitive dimension of CC competencies (Redman et

al., 2021), while also making efforts to address the skills dimension. This pattern may reflect greater confidence in assessing knowledge-related objectives. The participants' choices are consistent with the literature, which shows that competencies related to values, attitudes, and action appear to be addressed to a more limited extent (Monroe et al., 2019). Additionally, pre-service teachers placed particular emphasis on Systems Thinking in their assessment activities, recognizing it as a key CC competency and essential for CC teaching, in line with what is highlighted in the literature (Lorenzo-Rial et al., 2025; Levrini et al., 2019).

Moreover, the findings provide insights into how formative assessment elements were structured by pre-service teachers and how they were integrated into teaching scenarios. Participants placed emphasis on observing the overall learning process and planned to use, apart from common activities such as discussions and reports, a range of activities to elicit different types of information (e.g., written, oral, observational) about students' progress in CC competencies. This emphasis on "Eliciting Information" is in line with the literature, which highlights its importance for effective formative assessment (e.g., Dini et al., 2020; Ateh, 2015). In addition, feedback is either underutilized or avoided by pre-service teachers, highlighting a weakness in the integration of this element to their teaching (Ropohl & Rönnebeck, 2019). The lack of planning for providing and using feedback may have led to fewer planned formative assessment cycles and less support for learning goals, such as competency development (Cisterna et al., 2018). Therefore, pre-service teachers may benefit from support in feedback practices, as a core element of assessment to promote the development of competencies (Grangeat et al., 2021), and more specifically to foster climate literacy.

To conclude, this study highlights how pre-service teachers approach formative assessment for CC competencies at the planning level. The findings indicate that pre-service teachers place emphasis on addressing knowledge-oriented objectives and eliciting evidence of students' learning more systematically, while they struggle to plan systematic feedback and opportunities to act on it. Teacher education could equip pre-service teachers with the knowledge and skills needed to use contemporary assessment approaches, such as formative assessment.

Acknowledgement

The project is carried out within the context of the "PhD Candidates Scholarships" Programme - Category B, funded by the Special Account for Research Funds of University of Crete (Project's number: 12302).

References

- Annelin, A., & Boström, G. O. (2022). An assessment of key sustainability competencies: a review of scales and propositions for validation. *International Journal of Sustainability in Higher Education*, 24(9), 53-69. <https://doi.org/10.1108/IJSHE-05-2022-0166>
- Ateh, C. M. (2015). Science teachers' elicitation practices: Insights for formative assessment. *Educational Assessment*, 20(2), 112-131. <https://doi.org/10.1080/10627197.2015.1028619>
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in education: Principles, policy & practice*, 25(6), 551-575. <https://doi.org/10.1080/0969594X.2018.1441807>
- Bianchi, G., Pisiotis, U., & Cabrera, M. (2022). GreenComp. The European Sustainability Competence Framework. Publications Office of the European Union. <https://doi.org/10.2760/821058>
- Bryman, A. (2016). *Social research methods*. Oxford university press.
- Cebrian, G., Corres Gallardo, A. D. C., Renta Davids, A. I., & Junyent, M. (2025). Assessment of sustainability competencies in primary education: A systematic review. *The Journal of Environmental Education*, 56(3), 208-228. <https://doi.org/10.1080/00958964.2025.2471967>
- Cisterna, D., & Gotwals, A. W. (2018). Enactment of ongoing formative assessment: Challenges and opportunities for professional development and practice. *Journal of Science Teacher Education*, 29(3), 200-222. <https://doi.org/10.1080/1046560X.2018.1432227>
- Dini, V., Sevia, H., Caushi, K., & Orduña Picón, R. (2020). Characterizing the formative assessment enactment of experienced science teachers. *Science Education*, 104(2), 290-325.

<https://doi.org/10.1002/sce.21559>

- Grangeat, M., Harrison, C., & Dolin, J. (2021). Exploring assessment in STEM inquiry learning classrooms. *International Journal of Science Education*, 43(3), 345-361.
<https://doi.org/10.1080/09500693.2021.1903617>
- Grob, R., Holmeier, M., & Labudde, P. (2024). Analysing formal formative assessment activities in the context of inquiry at primary and upper secondary school in Switzerland. *International Journal of Science Education*, 1-22. <https://doi.org/10.1080/09500693.2019.1663453>
- Harris, C. J., Krajcik, J. S., Pellegrino, J. W., & DeBarger, A. H. (2019). Designing knowledge-in-use assessments to promote deeper learning. *Educational measurement: issues and practice*, 38(2), 53-67. <https://doi.org/10.1111/emip.12253>
- Haug, B. S., & Ødegaard, M. (2015). Formative assessment and teachers' sensitivity to student responses. *International Journal of Science Education*, 37(4), 629-654.
<https://doi.org/10.1080/09500693.2014.1003262>
- IPCC. (2023). Climate change 2023: Synthesis report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (H. Lee & J. Romero, Eds.). IPCC. <https://doi.org/10.59327/IPCC/AR6-9789291691647>
- Levrini, O., Tasquier, G., Branchetti, L., & Barelli, E. (2019). Developing future-scaffolding skills through science education. *International Journal of Science Education*, 41(18), 2647-2674.
<https://doi.org/10.1080/09500693.2019.1693080>
- Lorenzo-Rial, M. A., Varela-Losada, M., Pérez-Rodríguez, U., & Vega-Marcote, P. (2025). Developing systems thinking to address climate change. *International Journal of Sustainability in Higher Education*, 26(1), 83-100. <https://doi.org/10.1108/IJSHE-12-2022-0404>
- Monroe, M. C., Plate, R. R., Oxarart, A., Bowers, A., & Chaves, W. A. (2019). Identifying effective climate change education strategies: A systematic review of the research. *Environmental Education Research*, 25(6), 791-812. <https://doi.org/10.1080/13504622.2017.1360842>
- Redman, A., Wiek, A., & Barth, M. (2021). Current practice of assessing students' sustainability competencies: A review of tools. *Sustainability Science*, 16, 117-135.
<https://doi.org/10.1007/s11625-020-00855-1>
- Ropohl, M., & Rönnebeck, S. (2024). Making learning effective—quantity and quality of pre-service teachers' feedback. In *Developing Formative Assessment in STEM Classrooms* (pp. 63-83). Routledge. <https://doi.org/10.1080/09500693.2019.1663452>
- Taurinen, J., Vesterinen, V. M., Veijonaho, S., Siponen, J., Riuttanen, L., & Ruuskanen, T. (2024). Climate change competencies from perspective of Finnish youth. *Journal of Youth Studies*, 1-20.
<https://doi.org/10.1080/13676261.2024.2343724>
- United Nations Educational, Scientific and Cultural Organization (UNESCO). Education for Sustainable Development Goals. UNESCO: Paris, France, 2017.
- Wiek, A., Withycombe, L., & Redman, C. L. (2011). Key competencies in sustainability: a reference framework for academic program development. *Sustainability science*, 6, 203-218.
<https://doi.org/10.1007/s11625-011-0132-6>