IL NUOVO CIMENTO **46 C** (2023) 199
DOI 10.1393/ncc/i2023-23199-1

Colloquia: IPER 2022

An approach to research-based design of teaching-learning sequences in the context of physics education: Theoretical frameworks, pedagogical methods, and examples of Data Analysis

C. Fazio([1]), A. Agliolo Gallitto([1]), C. G. Galiano([1]), G. Giarratano([2]), I. Grazia([1]), G. Termini([1]) and O. R. Battaglia([1])

([1]) *Department of Physics and Chemistry "Emilio Segrè", University of Palermo - Palermo, Italy*
([2]) *Department of Psychological, Pedagogical, Physical Exercise and Training Sciences, University of Palermo - Palermo, Italy*

received 13 July 2023

Summary. — In this paper we discuss an example of the research approaches practiced at the University of Palermo Physics Education Research Group (UoP-PERG). Particularly, we discuss the design and development with groups of high school students of two Teaching-Learning Sequences (TLSs) on surface phenomena in liquids, and the data collection and analysis methods. In the introduction, we briefly discuss the educational reconstruction model, that we use as a theoretical framework for the designing of the TLSs, and a pedagogical methodology that in the last years has gained consensus among educators, *i.e.*, active learning. Some considerations on active learning pedagogical and cognitive psychology foundations are also made. The main aim of the TLSs is to improve students' learning of surface phenomena, a topic that is quite relevant in Physics and other scientific and technical fields. We provide a research-based, conceptual scheme of what we mean by "improvement of students' learning", and present the TLSs phases, that are based on the well-known inquiry- and investigation-based learning approaches, that are science-specific applications of the general idea of active learning. Then, we describe some methods we use in our researches to collect and give an example of the data analysis needed to study the progression of student learning, with respect to the conceptual scheme provided, and of some of the results obtained.

1. – Introduction

One of the main aims of science education research is to discuss the issues related to science teaching and student learning. Many theoretical frameworks [1-4] and methodological approaches aimed at improving the effectiveness of both teaching and learning have been proposed and validated in the past years. On the basis of these frameworks and methodological approaches, research focused of the design, implementation and validation of topic-oriented educational pathways aimed at helping teachers plan and monitor their instruction and, as a result, enhance students' learning. Examples found in the literature refer to pathways called Teaching-Learning Sequences (TLSs) (*e.g.*, [5, 6]) or Learning Progressions (LPs) (*e.g.*, [7-9]).

A characteristic of TLSs and LPs is a structure founded on research-based evolutionary processes aiming at making scientific and learner perspective interact, in order to optimize the learning outcomes and their transferability [5]. TLSs and LPs are typically carefully sequenced sets of learning blocks that students must master en route to mastering a further-reaching curricular objective. Frequently, these components include subskills and bodies of enabling knowledge. TLSs and LPs are usually planned and utilized as both research tools, to examine the efficacy of the proposed activities and of the evaluation tools, and as innovation proposals, to demonstrate how specific topic-related learning issues can be addressed.

Throughout the last few decades, numerous theoretical frameworks and instructional strategies have been proposed for the purpose of developing educational pathways in the field of science education and enhancing their design and implementation. Among these theoretical frameworks, the "educational Reconstruction" model developed by Kattmann *et al.* in 1995 [2-4] is one of the most widely used and discussed. It is predicated on the notion that an effective teaching-learning sequence should be based on both a close analysis of science content structure and educational issues, such as knowledge of the context, student common sense ideas, research-based learning knots and processes, and social and ethical implications. In order to "reconstruct" the content to be taught and construct the educational path to be tested with the students, it is necessary to take into account both the structure analysis of the science content and educational issues. The model is based on an integrated constructionist perspective: knowledge acquisition is viewed as a process in which the learner actively participates within a social and material context, and scientific knowledge is viewed as a tentative human construction [6].

As a result of this integrated constructionist perspective, the majority of teaching strategies used in the planning and development of TLSs and LPs are based on the general concept of "Active Learning". In fact, research in Science Education has shown that a change in teaching pedagogy from deductive-based approaches to approaches based on the active involvement of the learners in the construction of their knowledge can promote student learning. That is particularly documented with regard to the improvement of students' conceptual understanding, longer-term knowledge retention, and the development of critical thinking skills (*e.g.*, [10-12]). In "active learning", the learner does more than simply listen to a lecture. He/she is actively engaged in reading, writing, posing and discussing questions, gathering data from various sources, constructing models, and solving problems in an effort to construct and develop his/her knowledge, skills, and attitudes.

As it is clear, the ideas underlying active learning are based on the constructionist theory of learning, which describes how individuals can acquire knowledge and learn effectively. Learning is a dynamic process consisting of successive stages of adaptation to reality during which learners actively construct knowledge by developing, testing, and reframing their worldviews. The literature on constructionist models of human learning, so, suggests that useable knowledge is best gained in active learning environments (*e.g.*, [13]).

It is now acknowledged that planning educational activities with significant cognitive psychology research findings in mind can further enhance learning by positively impacting students' motivation, self-confidence, and intellectual development. Numerous proposed approaches to science learning, such as the well-known Investigative Science Learning

Environment (ISLE) approach [14], engage students in pedagogical activities that mirror scientific practice and cooperation, with the goal of fostering in them deep and meaningful learning, a "growth mindset" [15], and an overall sense of learning satisfaction.

According to American psychologist Carol Dweck, learning is highly dependent on whether students believe their abilities are fixed or can change. A "growth mindset" prepares students to actively and consciously direct their own learning through the activities proposed by the teacher. In this way, students are encouraged to put forth deliberate [16] and contextualized [17] effort and practice at increasing levels of complexity, with the goal of leaving the "zone of cognitive comfort" associated with things they feel already proficient at.

Despite of some criticism on the efficacy of Dweck's mindset approach to education, also related to teacher beliefs [18, 19], studies have demonstrated that mindset interventions can work in many contexts [20, 21]. When students believe they can improve, they exert effort in activities such as learning. Therefore, efforts, time, and support in completing tasks of increasing complexity enable students to acquire skills comparable to those of an expert, promote conscious and persistent learning, and foster the development of self-confidence and metacognition. An essential aspect of deliberate practice is the exercise and active development of skills at progressively higher levels, which enables students to acquire these skills most effectively [22]. Students can develop a personal awareness of their knowledge and skills through deliberate practice processes. This can enable them to better identify their strengths and weaknesses and reflect on how to optimize their learning. Students can also reflect on their personal learning/cognitive style(s) [23-26] and the level of their learning in a given context because of these procedures. The objective is to assist students in developing and empowering new skills and cognitive functions, while preventing them from spending too much time in their cognitive comfort zone, which, as stated, is frequently counterproductive to appreciable cognitive growth.

The action of reflecting on students' learning/cognitive style(s) is central for the teacher/researcher in planning and implementing a TLS/LP. In fact, an educational pathway aiming at being significant for the students should also take into account the way in which students prefer to acquire information from their environment (including the learning environment) and elaborate it in order to construct descriptions and explanations of real-world phenomena and situations. In other words, knowledge of theories regarding students' possible learning styles, such as the well-known VARK model [24,25], a refinement of the earlier VAK model of learning styles [23], can be useful for planning and implementing educational pathways that can be effective in promoting student learning, the ultimate goal of any educational action.

2. – The research

The main aim of this paper is to present an example of the research methods used at the University of Palermo Physics Education Research Group (UoP-PERG), discussing some typical pedagogical approaches used by the Group, and the methods used to collect and analyze data coming from the trialling of educational pathways. We will focus on some aspects of the design and trialling with groups of high school students of two TLSs on surface phenomena in liquids with respect to specific research variables related to learning, and we will give some results of the trialling with respect to those variables.

The surface phenomena content was selected because in physics, as well as other scientific and technical fields, a thorough comprehension of these phenomena is important. The pedagogical methods traditionally used to introduce students to the content are based on transmissive approaches dealing with macroscopic descriptions and, at higher educational levels, sometimes with an introduction to molecular interactions. However, since many years, research has demonstrated that these methods may not be entirely effective at capturing student interest (*e.g.*, [27]) and promoting students' meaningful comprehension of the physical content. For these reasons, we believe it is useful to suggest a new method for introducing students to the study of surface phenomena, based on the active participation of students in laboratory and modeling activities, as well as on the use of computer simulations for modelling purposes. Computer tools, such as computer-assisted data loggers and simulations, can be extremely beneficial for students, as they enable them to easily collect data in real-time and control parameters pertinent to understanding the mechanism of operation underlying the phenomena they wish to study, thereby largely fostering model-based reasoning (*e.g.*, [28]). Research demonstrates that models constructed at an intermediate scale (*i.e.*, mesoscopic scale) can be utilized effectively in science education. Particularly, the literature recognizes mesoscopic models as valuable for efficiently introducing topics such as solid friction and fluid statics [29]. These models have the benefits of the microscopic model. Particularly, they foster understanding based on the recognition of causal relationships [30] and "mechanism of functioning", a mental process that is at the basis of the development of explicative lines or reasoning. Furthermore, these models do not require a significant amount of computer resources to execute simulations implementing the models.

Starting from these considerations, we hypothesized that choosing an appropriate modelling scale to introduce the surface phenomena could appreciably enhance the teaching-learning processes, both at school and university levels.

We designed two TLSs, one based on macroscopic modelling, and the other based on mesoscopic modelling, which were pilot trialled each with a group of upper secondary school students. Each TLS was based on an inquiry/investigative-based approach and was planned to actively involve the students in constructing their learning. They were involved in posing questions, gathering information (through experiments, simulations, books, the internet, etc.), discussing and contrasting results in small and large groups, reaching consensus and sharing knowledge. It is worth noting that in the literature there are not many examples of researches aimed at studying the understanding of surface phenomena. For this reason, the main goal of the trialling was not to simply identify which group highlights the best learning depending on the different modelling approach, but to verify, on the basis of the abovementioned research hypothesis, the aspects of each approach that can be considered truly relevant in promoting learning. In that way, we are able to use data coming from the TLSs' trialling to create a new version of the TLSs that incorporates all the factors that have shown to be crucial for improving and promoting student learning. That new version is currently further trialed to extend the study with other Upper Secondary School students and will be updated for trialling at undergraduate level in the next academic year.

For the abovementioned reasons, we can state the main research question that guided us as:

Which parts of each TLS can be deemed important for promoting significant student learning?

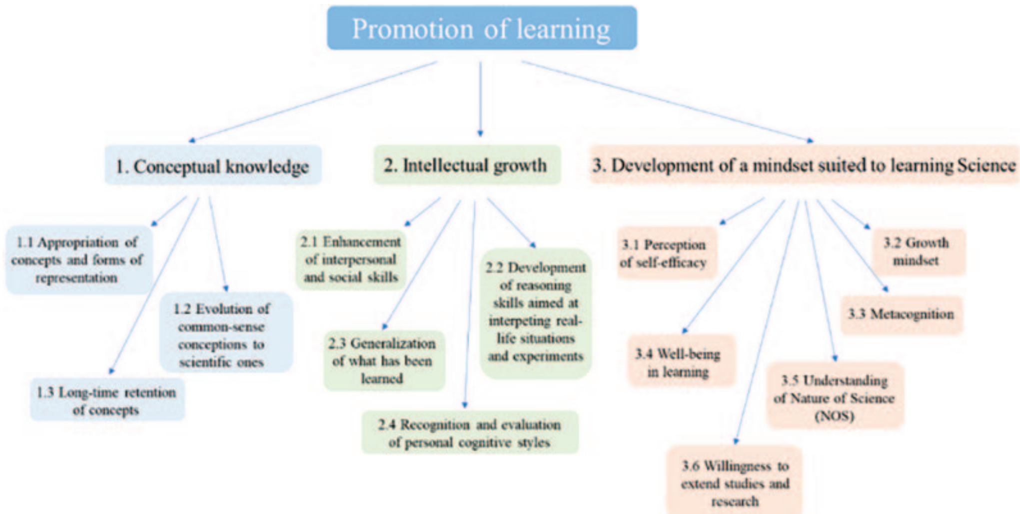

Fig. 1. – Scheme showing the research-informed dimensions related to promoting student learning of Science we chose in our research.

However, even a quick literature research shows that "promoting student learning" is a complex concept to study and assess. A comprehensive literature review of the "dimensions" (*e.g.*, [31]) that researchers consider when analyzing this concept led us to the construction of a conceptual scheme, reported in fig. 1. It depicts what we understand by "promotion of student learning" in our overall research. "Acquisition of conceptual knowledge", "Intellectual growth", and "Development of a mind-set suited to learning Science" are highlighted in fig. 1 as the three main research-informed dimensions that may help us to characterize the general concept of promotion of learning in a teaching-learning sequence, with specific reference to scientific learning. For each of these dimensions, the figure provides additional research-informed detail on the "sub-dimensions", henceforth referred to as "study variables", or "variables" for simplicity, that we aim to study throughout the research.

Each variable reported in fig. 1 has been identified through a thorough literature research, and should be carefully described to understand well its role in the study of the overarching concept of "promotion of learning". As in this paper we do not have enough space to do that, we report in table I some literature references that allow

2˙1. *The student sample.* – The TLSs were pilot tested with Upper Secondary School students during the Academic Year 2021–22 and we will briefly discuss their main phases below.

The sample on which the two TLSs were tested consists of forty pupils in their fourth year at "Liceo Scientifico", the Italian science-focused Upper Secondary School (age range 16–17, 20 females and 18 males). All the students have been studying physics since the first year of Liceo Scientifico. The research sample was composed of students coming from four different classrooms in the same school in Palermo, Italy. Both the physics teachers of these students hold master's degrees in physics. Moreover, they have a similar approach to teaching since they were all trained in professional development activities

Table I. – *Sub-dimensions of learning (variables) that we identified for our study, and some related literature references.*

N	Variable	Literature references
1.1	Appropriation of concepts and forms of representation	[32-34]
1.2	Evolution of common-sense conceptions to scientific ones	[35, 36]
1.3	Long-time retention of concepts	[37-39]
2.1	Enhancement of interpersonal and social skills	[40, 41]
2.2	Development of reasoning skills aimed at interpreting real-life situations and experiments	[30, 42, 43]
2.3	Generalization of what has been learned	[44-46]
2.4	Recognition and evaluation of personal cognitive skills	[23-26, 47]
3.1	Perception of self-efficacy	[48-51]
3.2	Growth mindset	[15, 16, 52]
3.3	Metacognition	[53, 54]
3.4	Well-being in learning	[55-58]
3.5	Understanding of Nature of Science	[59, 60]
3.6	Willingness to extend studies and research	[61, 62]

at the Università degli Studi di Palermo, which prepared them for the use of inquiry methodologies. Therefore, the sample of students was selected to be as homogeneous as possible in terms of preparation, method of study, and motivation.

Students voluntarily choose to participate in the trialling, during the afternoons, after their regular school lessons. Due to the small size of the student sample (due to space and equipment constraints in our didactic laboratory), we randomly picked students from different courses in order to obtain two 20 student subsamples. Moreover, as collaboration and group-work are integral elements of both the TLSs, creating working groups comprised of students who do not know each other ensures that all students may find their place in the group and freely express themselves away from the dynamics of the class from which they came. For instance, the shyest pupils may hide behind the most outgoing classmates, while the most intelligent students could "leap over" the more insecure students. The first subsample, "Group A", investigated surface phenomena using a conventional macroscopic modelling approach. Group B, the second subsample, investigated the identical issues as Group A from a mesoscopic modelling viewpoint.

2·2. *The TLSs.* – As we said, both TLSs are inspired by the ISLE approach to science education. Particularly, they involve observation, experiments, modeling, small and large group discussions, and pay attention not only to the cognitive aspects of learning, but also to the affective ones, that are relevant in fostering effective and persistent learning. According to this approach, to generate interest in learning, students must be authentically involved in pedagogical activities, and specifically in science practices simulating scientific research. Particularly, they must alternate between working in small and large groups. When working in small groups, they observe natural phenomena, pose questions related to what they observe, plan experiments and perform them, formulate answers and multiple explanations, and test them systematically on the experimental data. When discussing in a large group, students collaborate and share their ideas, continuously evaluating and enhancing the models created in small groups, and comparing and contrasting them within a model of a scientific community created in class.

During the planning and implementation of the TLS, students' teamwork phases and approach to physical situations and the descriptions/explanations were given special consideration. Not only were students encouraged to collaborate with their teammates, but they were also urged to personally reflect on their work and on the agreement with the group conclusions. In this way, students could independently search for sources and materials that assisted them in carrying out the activities and consolidating the acquired knowledge, allowing them to discuss the results with their teammates in a proactive manner.

During small and large group activities, students actively planned and conducted observations and experiments, and constructed descriptive and explanatory models of the observed phenomenology. The instructors paid special attention to activating gradually increasing levels of complexity and difficulty in the proposed tasks (*e.g.*, [16]). In addition, we took special care in our approach to propose activities that could accommodate the various learning and cognitive styles that students may possess (*e.g.*, [23-26]) in order to optimize the perception, collection, and processing of learning materials [47]. At the conclusion of the activities, we asked each group to compile a report on their scientific experience using notes, photographs, videos, and any other data they had collected.

During the first lesson of the TLS, students were given a pre-instruction survey on fundamental issues related to surface phenomena that were not directly related to the topics presented and analyzed throughout the course. After a thirty-minute break on the same day, students were given a second questionnaire regarding the teaching-learning sequence. Both questionnaires were designed and validated using well-known methods described in the literature [63]. The questionnaires were validated using samples distinct from those used to test the teaching-learning sequences, and comprised of fourth-year students from the same school as the sample used for the study.

Observation of phenomena through audio/video recordings and the gathering of additional information through various media (books, journals, web videos, websites, etc.) were performed with increasing levels of complexity and student participation in the classroom. The experimental, modeling, and simulation activities were also carried out with progressively increasing levels of complexity, and information was conveyed and shared utilizing a variety of communication styles. During the activities, students were first asked to make predictions and compare them to the experimental results; they were then invited to discuss and express their agreement with the group's conclusion on an individual basis. Students were always asked to consider their perceptions of self-efficacy in understanding and well-being in learning.

Every student participated in both qualitative and quantitative laboratory activity. Following the bounded inquiry approach [64] qualitative laboratory activities were conducted in which students were free to choose which tools to use and which path to follow in order to answer the questions posed by the instructors.

In addition to qualitative experiments, useful for introducing students to the situations to be analyzed, we designed quantitative experiments to investigate surface phenomena and provide estimates of surface tension values in various liquids. The students were guided, by means of a guided inquiry approach [64,65], through the quantitative activities, with the aim to lead them to answer the questions posed by the instructors.

In some instances, beginning with well-known experimental setups, we reconstructed them on a shoestring budget using materials readily available in standard didactic laboratories. Students were able to measure surface tension values for various liquids, including commercial demineralized water, 99% ethyl alcohol, pure glycerol, common commercial peanut, sunflower, and corn oil, and a 30 ml mixture of dishwashing soap and demineralized water using the Du Nouy ring method. The ring was made of aluminium, and the forces involved were measured using a digital balance (sensitivity: 10^{-4} N) [66]. In another experiment, students were able to estimate the surface tension of soap bubbles as the radius of the bubbles varied so as to verify the Young-Laplace law (For more detail see [67]). Finally, in a third experiment, students were able to estimate the contact angle between a liquid such as water and glass [66].

The modelling phases of the TLS trialled by Group A are developed from a macroscopic perspective, by introducing students to the general concept of surface tension and introducing cohesive and adhesive forces on liquid and solid surfaces as the cause of surface phenomena. Students studied books and watched videos of lessons and simulations (*e.g.*, [68-73]), drawing and discussing graphical schemes of forces acting on liquid and solid surfaces on the basis of what is reported in the books/videos inspected.

The modelling in the TLS trialled by Group B is based on mesoscopic modelling of surface tension [74, 75], in which the liquid is described as a set of Lagrangian particles and the interactions from which surface tension of the liquid arises are described in terms of attractive and repulsive forces among particles. The equations of motion (Navier-Stokes equations) of the discretized liquid are solved through the Smoothed-Particle Hydrodynamics (SPH) computational method [76]. Group B students were introduced to the model without discussion of the SPH method's mathematical specifics and focused on the different roles played by forces over small and large distances, as well as the different interactions between "liquid" and "solid" particles. They were allowed to modify some simulation parameters and see in real time the effects on the simulations.

At the conclusion of the activities, students answered two post-instruction questionnaires identical to the pre-instruction ones. After instruction, a satisfaction survey regarding the TLS activities and methodologies was also administered to the students. A couple of months after the conclusion of the activities, students responded to a questionnaire identical to the post-instruction one which focused on teaching-learning specific topics.

Table II resumes the timeline for implementation of the TLSs and provide a brief description of activities carried out by students in Group A and B.

2·3. *Data collection.* – The data were collected by means of different instruments. In educational research, the choice of data collection instruments is strongly based on the type of analysis that is to be conducted in order to address the specific research problem. Considering the complexity of the concept we want to study ("promoting student learning"), we used a variety of instruments, resumed as follows:

- A questionnaire (Q1) on general topics, related to surface phenomena but not specifically related to the topics discussed and analyzed during the TLSs.
- A questionnaire (Q2) on topics specific to the TLSs.
- A questionnaire (Q3) to investigate students' satisfaction about the TLSs.
- Students' worksheets.

TABLE II. – *Sub-dimensions of learning (variables) that we identified for our study, and some related literature references.*

	Group A	Group B
Day 1	Administration of Pre-Instruction questionnaires Q1 and Q2	Administration of Pre-Instruction questionnaires Q1 and Q2
Day 2	Qualitative activities: Observation: gerridae on the water surface Qualitative experiment: Objects on the water surface Qualitative experiment: soap water films in metal frames	Qualitative activities: Observation: gerridae on the water surface Qualitative experiment: Objects on the water surface Qualitative experiment: soap water films in metal frames
day 3	Qualitative activities and conceptual pit stop: Qualitative experiment: liquids in capillary tubes Qualitative experiment: sessile drops and contact angles Qualitative experiment: objects on the water surface when soap is added Conceptual pit stop	Qualitative activities and conceptual pit stop: Qualitative experiment: liquids in capillary tubes Qualitative experiment: sessile drops and contact angles Qualitative experiment: objects on the water surface when soap is added Conceptual pit stop.
day 4	Macroscopic modelling: The researchers provide students with an explanation of surface phenomena based on the macroscopic model, as found in most textbooks. To support a description of surface phenomena based on this approach, researchers use videos, images and diagrams strongly based on a a macroscopic view of surface phenomena.	Mesoscopic modelling: Formation of a liquid drop in absence of gravity; formation of liquid menisci; formation of a liquid sessile drop on a solid surface.
day 5	Quantitative experiments: Measurement of the surface tension of water by the ring method; measurement of the surface tension of water by the ring method; measurement of the surface tension of water by the water drop method; measurement of the water-glass contact by the variable section capillary method.	Quantitative experiments: Measurement of the surface tension of water by the ring method; Measurement of the surface tension of water by the ring method; measurement of the surface tension of water by the water drop method; measurement of the water-glass contact by the variable section capillary method.
day 6	Final discussion Administration of Post-Instruction questionnaires Q1 and Q2. Administration of satisfaction questionnaire Q3.	Final discussion Administration of Post-Instruction questionnaires Q1 and Q2. Administration of satisfaction questionnaire Q3.

- Audio recordings of students group discussions at the end of each activity.
- Students' feedback on the activities carried out during the TLSs.
- Students' contributions during the final day brainstorming phase.
- Notes from researchers.

All these instruments allowed us to build different databases, that were analyzed by means of different methodologies. In general, from each database collected with a specific instrument it is possible to extract information on some of the dimensions of learning that we have chosen to study. However, as we will see, some databases are more suitable than others for investigating a given dimension/variable related to learning. All the instruments used to collect the data are well-known in the literature and commonly used in research in education.

Questionnaires Q1 and Q2 were administered before and after instruction (pre-instruction Q1 – post- instruction Q1 and pre-instruction Q2 – post-instruction Q2), with the aim to get information on some of the study variables related to the first two dimensions of learning that we identified, "Acquisition of conceptual knowledge", and "Intellectual growth". Particularly, questionnaire Q1 was mainly aimed at obtaining information on variables 1.2: Evolution of common-sense conceptions to scientific ones, and 2.2: Development of reasoning skills aimed at interpreting real-life situations and experiments. Questionnaire Q2 was aimed at obtaining information on variables 1.1: Appropriation of concepts and forms of representation, 2.2: Development of reasoning skills aimed at interpreting real-life situations and experiments, and 2.3: Generalization of what has been learned. Questionnaire Q2 was also administered two months after the end of the pedagogical activities (post-post-instruction Q2), in order to get information also on the study variable 1.3 (Long-time retention of concepts).

Questionnaire Q3 was aimed at studying variables related to the third dimension of learning, "Development of a mindset suited to learning Science", and was administered only at the end of the pedagogical activities. It is a Likert scale questionnaire which also includes a final open-ended question in which students are asked for sincere feedback on their experience during the experimentation.

All the other data collection instruments were aimed at completing the study, giving detail on almost all the study variables (except variable 1.3, that was studied only by means of post-post-instruction Q2). At the end of each day of experimentation, the researchers noted their considerations in a logbook. These notes allowed us to check for consistency of what was found in the other databases and sometimes to reconstruct the audio recordings when they were not clear.

2˙4. *Questionnaire validation*. – Questionnaires Q1, Q2, and Q3 have been in part designed by the researchers, by using literature sources. They were all validated according to methods well known in the literature [63]. We briefly describe the validation procedures we used as follows:

- Content/logical validation. It is a kind of validation procedure aimed at allowing researchers to understand how the test items are representative of the content they aim to investigate on. The reliability of this validation is influenced by content experts' judgment, since they are designated to indicate whether the test is suitable to measure what it aims to measure. All the researchers participated to content validation of the questionnaires.

- Construct validity. It allows researchers to understand if a test reflects or not a given construct. Constructs are abstractions used by the researchers to conceptualize a latent variable which is correlated with scores on a given measure, although it is not directly observable. All the researchers participated to construct validation of the questionnaires.

- Face validity. It is a fundamental step to measuring the validity of a test, as it studies how the questionnaire questions are understood by students, and allows the researcher to modify "on the fly" a question that is not clear to students and verify the relevant effect. Questionnaires Q1, Q2, and Q3 were face-validated with a sample of 15 students attending the same school asof the students of the research sample, at different classes of the same grade.

2˙5. *Data analysis*. – Our research embraces different kinds of data collection instruments and analysis techniques, and can, therefore, be defined as a research based on mixed-method approaches. Such approaches turn to be particularly dynamic and suitable to expand research aims and improve the analytic power of studies.

All the data we collected were coded and analyzed by means of different methods, depending on their nature. In this way, we extrapolated detailed information and insights on the data that allowed us to achieve a meaningful and deepen interpretation of them, in the light of our research aims.

Data coming from questionnaires Q1 and Q2 were studied by means of phenomenographic methods [77] and refined by means of content analysis methods [78]. Data were coded in terms of the answers most frequently given by the students to the questions [79, 80], and quantitatively treated as briefly described below.

Questionnaire Q3 was a Likert scale questionnaire, with a final open-ended question in which students were asked for their sincere feedback on their experience during the experimentation. The scale was converted in a number scale from 0 (not at all) to 4 (very much). For each question, the weighted mean value of the scores assigned by the students $\overline{x} = \sum_{i=1}^{N} \frac{x_i N_i}{N}$, where N_i is the number of students who assigned the score x_i, and N in the total number of students, was determined.

Data coming from the other sources (students' worksheets, audio recordings of students group discussions at the end of each activity, students' contributions during the final day brainstorming phase, and students' feedback on the activities carried out during the TLSs) were analyzed with methods inspired by thematic analysis [81, 82], so to synthesize their richness and complexity.

2˙5.1. Analysis of answers to questionnaires Q1 and Q2. The analysis of questionnaires Q1 and Q2 was performed on the basis of the variables we wanted to obtain information about by means of these instruments (variables 1.1 (Q2), 1.2 (Q1), 1.3 (Q2), 2.2 (Q1 and Q2), and 2.3 (Q2)). For each of these variables, we first identified levels related to students understanding of the themes related to them. For example, regarding variables 1.2: Evolution of common-sense conceptions to scientific ones, and 2.2: Development of reasoning skills aimed at interpreting real-life situations and experiments, we performed an analysis of student answers to questionnaire Q1 (both variables) and Q2 (only variable 2.2) aimed at highlight how, before instruction, students describe and explain surface phenomena, and if, and how, these descriptions/explanations are modified when attending the course.

As we said above, the analysis of students' answers to the questionnaire was performed on the basis of phenomenographic and content analysis methods. It allowed us to identify,

Table III. – *Students' "epistemological profiles" related to the ways of tackling the questionnaire and the related reasoning procedures.*

Practical/Everyday	Descriptive	Explicative
Reflects the creation of situational meanings derived from everyday contexts. The student uses other situations, perceived as analogous to the one proposed in the question, to try to describe/explain it.	The student describes and characterizes the proposed situation/analyzed process by searching in memory the variables perceived as relevant and/or recalling their relations. The variables and the relationships among them are expressed by means of different languages (verbal, iconic, mathematic). Causal relations among the variables on the basis of a functioning model (microscopic/macroscopic) are not given.	The student gives an explanation of the proposed situation referring to a model qualitative and/or quantitative) based on cause/effect relations. They may also provide explanatory hypotheses by introducing models which can be seen at a theoretical level.

also based on previous research [80, 81, 83, 84], three students' "epistemological profiles", related to three different ways of reasoning when tackling the situations proposed in the questions.

The profiles are resumed in table II, where a brief description of the reasoning procedures that the students used when tackling the questions is given. We note that the Practical/Everyday profile can be related to the use of common-sense knowledge. Both the Descriptive and Explicative profiles are related to the use of scientific knowledge, although at different levels of sophistication, as it is evident from table III. For this reason, the analysis performed by means of the individuations of the three abovementioned profiles gave us insights on the evolution of both study variables 1.2 and 2.2.

Table IV shows some examples of keywords and sentences identified on the basis of the

Table IV. – *Examples of terms and sentences in students' answers used to classify them in one of the three "epistemological profiles" described in table II.*

Practical/Everyday	Descriptive	Explicative
(according to my) experience ...	I remember that ...	Molecular Movement ....
Like I see in real life ...	I studied that ...	... is similar to ...
Usually ...	I know that ...	Microscopic ...
Real object ...	The formula says ...	Inter-Molecular forces ...
Like an insect on water ...	The graph shows ...	Interaction ...
	There are adhesive and cohesive forces ...	Equilibrium ...
	There is surface tension ...	Molecules ...
	Chemistry/Physics say ...	

content analysis of students' answers, which allowed us to trace back students' response strategies to one of the three abovementioned profiles.

After the identification of the three epistemological profiles shown in table III, the answers of students to questionnaires Q1 and Q2 are classified into one of these profiles, by means of a several-stage negotiation among the researchers. Tables and graphs of the coding allows the researchers to study the related variables (1.2 and 2.2), *i.e.*, how students' answers move from one profile to another as a consequence of the TLS's attendance.

A deeper analysis of the answers given to questionnaires can be performed by means of Cluster Analysis [84]. Many disciplines of research, including information technology, biology, medicine, archeology, econophysics, and market research, employ ClA techniques [85-88]. These techniques enable the researcher to identify subsets or clusters of objects that tend to be homogeneous "in some sense" (unsupervised classification [89,90]) without having prior knowledge of the forms that these groups take. In accordance with the selected criteria, the results of the analysis can disclose a high degree of homogeneity within each group (intra-cluster) and a high degree of heterogeneity between groups (inter-cluster).

In our research group, we use ClA techniques in order to identify groups (*i.e.*, clusters), of students showing similarity in the ways they answer the questionnaire (*e.g.*, see [79, 83, 91]). This procedure can assist us in describing the behavior of students in relation to variables describing the learning dimensions depicted in fig. 1. Specifically, ClA techniques enable us to identify the characteristics that distinguish the responses of students belonging to the same cluster, as well as the differences and similarities between the responses of students belonging to different clusters. Thus, it is possible to infer typical "behavior profiles" of students with regard to a particular study variable, such as those presented in table I. For instance, a ClA enables us to characterize the sample in terms of the typical common-sense or scientific conceptions and the typical lines of reasoning utilized by students when responding to a questionnaire (*e.g.*, with reference, but not limited to, the epistemological profiles reported in table III), as well as the most common generalization skills or forms of content representations highlighted by students.

2˙5.2. Analysis of the other data. The analysis carried out on the data collected by means of the other instruments (students' worksheets, audio recordings of students group discussions at the end of each activity, students' feedback on the activities carried out during the TLSs, students' contributions during the final day brainstorming phase, and notes from researchers) was inspired by thematic analysis methods [81, 82], and involved the following steps:

1) Repeated readings of the data in order to become familiar with them.

2) Identification of text segments useful for answering the research question.

3) Identification of codes that synthesize the information conveyed by data.

4) Labelling of text segments of analytic interest (see step 2) through the codes identified in step 3.

5) Construction of a table of code-variable correspondences.

6) Identification of text segments significative for the analysis of specific aspects of learning (the study variables (or sub-dimensions of learning) introduced in fig. 1) based on the code-variable correspondences.

The identification of codes allowed us to synthesize the data complexity and to highlight how the variables, that give us information on the specific dimensions of learning, emerge from data. Labelling data segments by using codes allowed us to navigate more easily and intentionally within the datasets. We used an inductive coding. This means that we did not have a set codebook, but we created codes based on the qualitative data itself. Codes have been identified by us on the basis of the information carried out by data. In particular, the initial codes were roughly identified from the recurrences of words and sentences.

The codes identified after first reading the databases are reported below. For each code, we give a brief description clarifying which aspects of data can be summarized by the code.

Procedures-methodologies understanding. Students report on activities, tools, situations, etc. which promote or hamper the comprehension of specific topics addressed during the trialling stage.

Tools-skills. Students describe tools, materials, and skills acquired and/or used during the trialling stage.

Theory vs. practice. Students point out the difference between theory and practice. Theory is what they are most used to, practice is something they are not yet familiar with.

Traditional lecture vs. "innovative lecture". Students strongly perceive the difference between the traditional lectures they are used to at school and what they defined "innovative lectures" based on approaches they are not familiar with. By traditional lectures they refer to frontal lessons in which teachers explain and students listen to.

Content understanding. Different levels of understanding of the topics emerge from data. From students' answers to content questions, it emerges whether they have understood a content and/or its forms of representation consciously.

Debate. Students openly declare that the debate turns out to be an important tool to achieve a greater and better understanding of the topics.

Perspective. A given topic can be analysed from multiple perspectives and points of view. Students expand their learning perspective through the point of view of others, through the use of new study methods and new learning tools.

Role. Different roles can be assumed by students in their learning process. For example, students can distinguish the learning contexts in which they have an active or a passive role.

Language. The use of specific lexicon, that is scientific terminology used in a conscious way, is highlighted. Students recognize the role of mathematical language in the formalization of results obtained through the experiments. The acquisition and use of a scientific vocabulary facilitates the communication of the results among students.

Reflection. Students accurately reflect on the activities carried out, on what they have learned, on the skills they have acquired. They critically discuss the pros and cons of the activities they were involved in and reflect on how to apply models acquired in a given context to different situations.

Acknowledgement. Students declare they have acquired self-awareness by carrying out the activities proposed during the trialling stage. Students show to be aware of their strengths and weaknesses and acknowledge their progress.

Engagement-interest. Students talk about the activities they find most interesting and engaging. Many of them find computer-based simulations and hands-on experiments particularly challenging, while they get bored during more "traditional" activities similar to that they are used to at school.

Comfort. Students talk about the contexts in which they find themselves more comfortable during the learning process and those in which they did not feel comfortable.

Proactiveness. Students give us suggestions how to modify and improve didactic activities, based on what they have experienced during the trialling stage. Some of them reproduce or propose to reproduce some of the experiments carried out by making changes and look for additional information on the topics addressed in the classroom.

After repeated readings, we agreed that some codes could be merged with each other since they carry the same or similar information about the data. The codes "Procedures-Methodology understanding", "Tools-skills", "Theory *vs.* practice" and "Traditional lectures *vs.* innovative lectures" have been embedded into the overarching code A, the codes "Debate", "Perspective" and "Role" into C, the codes "Engagement-interest", "Comfort" and "Proactiveness" into F and finally, the codes "Reflection" and "Acknowledgement" into E.

Table V is a code-variable correspondence table. It shows which variables, and therefore which aspects of learning, emerge from the data labelled with a given code. For example, as can be seen in fig. 2 and in table VI, data labelled with the code "A" will hold information about aspects of learning related to intellectual growth (variables 2.2, 2.3) and development of a mindset suited to science (variables 3.2, 3.3, 3.5, 3.6). As can be seen from the table, each code can summarize information relating to one or more aspects of learning (*i.e.*, Conceptual knowledge, Intellectual growth, Development of the "right mindset"). The overarching codes (A-F) reported in table III are the codes that have been definitely used to conduct the qualitative analysis ofdata. These codes incorporate and synthetise all the information carried by the codes that have merged to constitute them.

Variable 1.3 is not present in table V. In fact, we could obtain information on it only by means of the post-post-instruction administration of questionnaire Q2. It is worth noting that table V is the result of a long negotiation process, which led to an agreement among the researchers involved in data analysis. In other words, the choice of "merging" two or more codes and the association of each code to one or more variables were discussed and agreed by the researchers.

3. – Some results of the data analysis

In order to give some details of the results we obtained and the way we formulated an answer to the research problem (on the basis of the study variables we identified for the research), we report here examples of the analysis of the data we collected during

TABLE V. – *Code-variable correspondence table showing what kind of information about the learning process is synthesized by each code. Codes composing each overarching code are reported.*

	Conceptual knowledge		Intellectual growth				Intellectual growth					
codes	1.1	1.2	2.1	2.2	2.3	2.4	3.1	3.2	3.3	3.4	3.5	3.6
A				X	X			X	X		X	X
B	X	X		X	X		X	X	X	X	X	X
C		X	X	X		X	X	X	X	X	X	X
D	X	X	X		X				X		X	
E				X	X	X	X	X	X			X
F								X		X		X

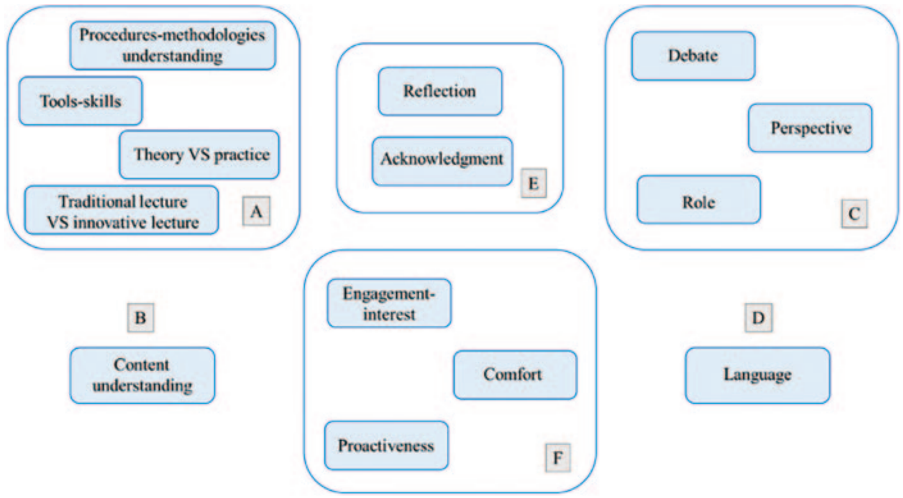

Fig. 2. – Diagram showing the codes introduced in the data analysis. The boxes contain the codes subtending to the same overarching code (A–F).

the TLS trialling with respect to some of the study variables. Particularly, we report an analysis of answers to questionnaires Q1 and/or Q2 questions with respect to variable 1.1: Appropriation of concepts and forms of representation (comparison between pre-, post- and post-post-administration of Q2), variable 1.2: Evolution of common-sense conceptions to scientific ones (comparison between pre- and post-administration of Q1), variable 1.3: Long-time persistence of concepts (comparison between pre-, post- and post-post-administration of Q2), and variable 2.3: Generalization of what has been learned (comparison between pre-, post-, and post-post administration of Q2). Moreover, an example of the results of textual analysis of the audio recordings of students' contributions during the final day brainstorming phase, with respect to all the study variables (according to table V) is reported.

3˙1. *Variables 1.1 and 1.3.* – This part of the study of variables 1.1: Appropriation of concepts and forms of representation, and 1.3: Long-time retention of concepts, was done by means of a content analysis of the answers to questionnaire Q2 before, after instruction and after the two-months break.

A further study of variable 1.3 was conducted on other databases and an example will be given in sect. 3.4.

Table VI. – *Codes composing each overarching code.*

A	Procedures-methodologies understanding/Tools-skills/Theory *vs.* practice/ Traditional lecture *vs.* innovative
B	Content understanding
C	Debate/Perspective/Role
D	Language
E	Reflection/Acknowledgment
F	Engagement-interest/Comfort/Proactiveness

TABLE VII. – $\chi^2 = 248.0$, $p < 1\%$.

Group A – Q2	Application	Correct description	Incorrect description	No answer
Pre-instruction	2	35	56	137
Post-instruction	40	98	22	9
Post-post instruction	30	110	14	15

TABLE VIII. – $\chi^2 = 174.4.0$, $p < 1\%$.

Group B – Q2	Application	Correct description	Incorrect description	No answer
Pre-instruction	2	65	39	82
Post-instruction	61	96	27	2
Post-post instruction	58	92	26	15

As the variable regards the appropriation of both concepts and forms of representations, we decided to analyse students' answers separately with respect to these two aspects.

3˙1.1. Concepts. References [32-34] discuss ways to define and evaluate the appropriation of concepts and content. According to these authors, teachers can observe a first level of appropriation when personal signature ideas grounded in the discipline emerge in students, and they are able to correctly describe and discuss contents. A second, deeper appropriation occurs when students are able to apply contents to solve problems and face situations.

For this reason, we decided to analyse the answers given by students to questionnaire Q2 searching in each student answer for the evidence of 1) application of the concept to solve a problem or face a situation; 2) correct description of a concept; 3) incorrect description of a concept; 4) no answer.

Tables VII and VIII show the contingency tables for the results of that analysis for Group A and Group B students, respectively, in the three administration phases of the questionnaire, and fig. 3 reports the related bar diagrams. The values represent the number of answers that highlight a specific level of appropriation of concepts.

As can be seen in fig. 3, in the pre-instruction questionnaire, the percentage of application of concepts is significantly low for both groups, and the percentage of no answers is particularly high, especially for Group A. From the pre- to the post-instruction questionnaire administration, an increase in the percentage of application of concepts and correct descriptions is registered, especially in Group B. These results are confirmed in the post-post instruction administration of the questionnaire.

3˙1.2. Forms of representation. For what regards the forms of representations, references [32-34] identify verbal, iconic, tabular, graphic, analytical representations as the ones most commonly used in science. So, we analysed the answers given to Q2 searching for the evidence of such kind of communication and representation channels in each student answer.

Tables IX and X show the contingency tables for the results of the analysis for Group A and Group B students, respectively, in the three administration phases of the questionnaire, and fig. 4 reports the related bar diagrams. The values represent the number of

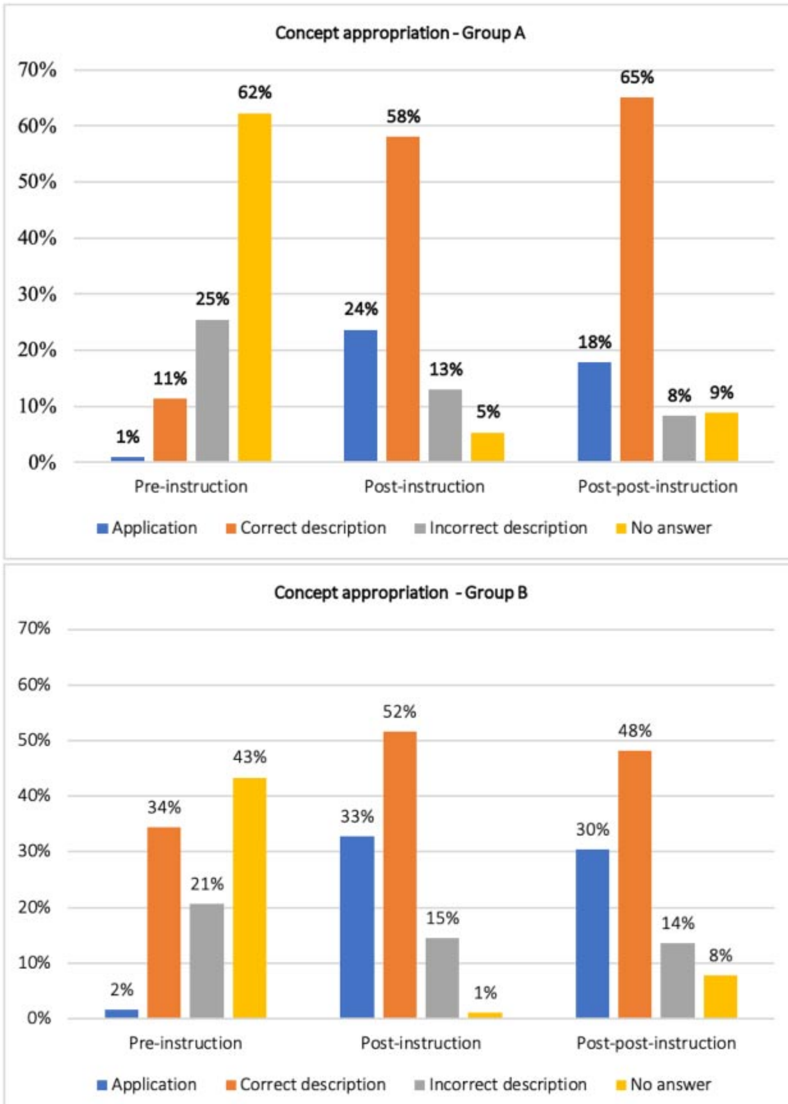

Fig. 3. – Bar diagrams showing the percentages of answers given by Group A and Group B students, respectively, to questionnaire Q2 during the pre-, post-, and post-post-instruction administrations. The answers are categorized according to the different level of concept appropriation identified in our analysis.

TABLE IX. – $\chi^2 = 233.4.0$, $p < 1\%$.

Group A – Q2	Verbal	Iconic	Tabular	Graphical	Analytical	No representation
Pre-instruction	43	15	10	13	2	137
Post-instruction	30	50	8	32	40	9
Post-post instruction	45	35	8	31	35	15

TABLE X. – $\chi^2 = 180.5$, $p < 1\%$.

Group B – Q2	Verbal	Iconic	Tabular	Graphical	Analytical	No representation
Pre-instruction	50	25	12	14	6	82
Post-instruction	30	71	4	59	20	2
Post-post instruction	35	68	8	50	15	15

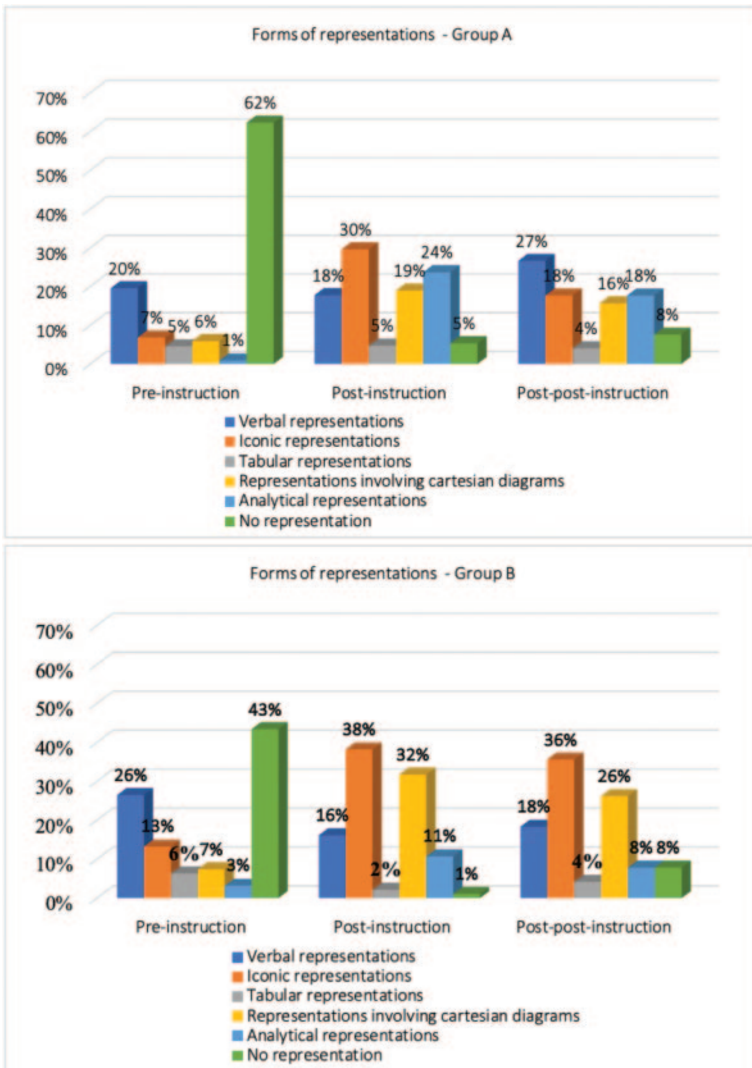

Fig. 4. – Bar diagrams showing the percentages of answers given by Group A and Group B students, respectively, to questionnaire Q2 during the pre-, post-, and post-post-instruction administrations. The answers are categorized according to the different forms of representation used by students.

TABLE XI. – $\chi^2 = 24.69$, $p < 1\%$.

Group A – Q2	Everyday	Descriptive	Explicative	No answer
Pre-instruction	58	45	41	31
Post-instruction	35	72	55	10

TABLE XII. – $\chi^2 = 55.07$, $p < 1\%$.

Group A – Q2	Everyday	Descriptive	Explicative	No answer
Pre-instruction	74	69	28	16
Post-instruction	26	64	80	4

answers that highlight the use of one representation channel.

As we can see in fig. 4, in answering the pre-instruction questionnaire, the prevailing form of representation is the verbal one, for both Group A and Group B. In answering the post-instruction questionnaire, Group A student show an increasing use of formulas (analytical representations) to address the proposed situations, while Group B students are more oriented to the use of schemes and graphs (iconic representations and representations involving Cartesian diagrams). These trends are confirmed in the answers of the post-post instruction questionnaire. In both groups, the percentage of students using tabular forms of representation is low and almost unchanged both in pre- and post-instruction administration of the questionnaire.

3·2. *Variable 1.2*. – The study of variable 1.2: Evolution of common-sense conceptions to scientific ones, was done by means of a phenomenographic and content analysis of the answers to questionnaire Q1 before and after instruction, according to what is reported in sect. 2.5.1 (tables II and III). A further study of the variable was conducted on Q2 answers, but we will not discuss it here.

Tables XI and XII report the contingency tables for the answers given to Q1 by Group A and Group B students, respectively, before and after instruction. In these tables, the number of answers is reported, and in fig. 5 the percentage of answers is reported.

Figure 5 gives detail about what happened in each group with respect to variable 1.2 in answering Q1. It is clear that in the post-instruction questionnaire, both groups still give everyday-type answers, but to a lesser extent than before. Both groups also show a decrease in the percentage of not answered questions from the pre- to post-instruction questionnaire administration. Group A highlights an increase of 16% in descriptive-type answers. Moreover, a decrease of 12% in the number of not-answered questions from the pre- to post-instruction questionnaire is highlighted.

On the other hand, Group B maintains the percentage of descriptive-type answers and highlights a significant increase (31%) in explicative-type answers. Everyday-type answers show a clear decrease (25%) from the pre- to post-instruction questionnaire.

3·3. *Variable 2.3*. – This part of the study of variable 2.3: Generalization of what has been learned, was done by means of the analysis of the answers to questionnaire 2 before, after instruction and after the two-months break. A further study of the variables was conducted on the other databases and an example will be given in sect. 3.4.

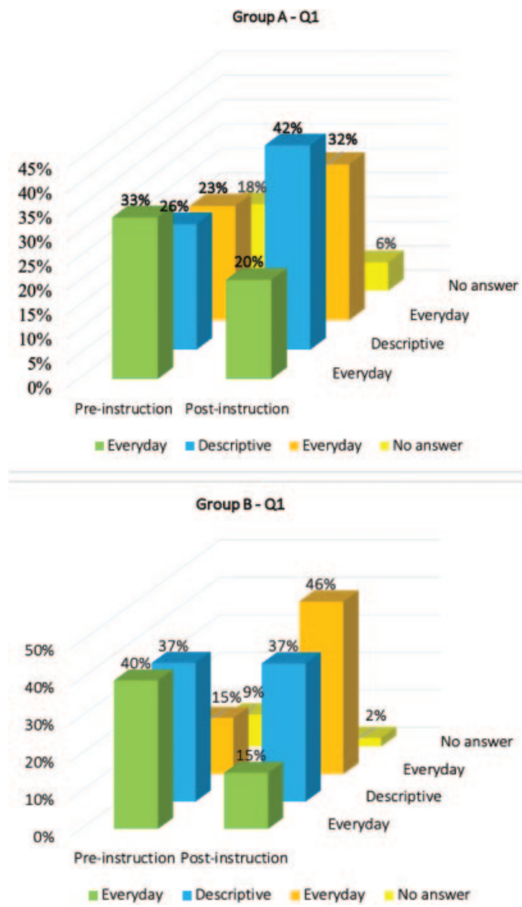

Fig. 5. – Bar-diagrams showing the percentages of answers given by Group A and Group B students, respectively, to the pre-and post-instruction questionnaire Q1. Answers are categorized according to the "epistemological profiles" identified in our analysis and reported in table III.

References [44-46] show that a way to discuss ways to evaluate variable 2.3: Generalization of what has been learned, is to study the use of contents and techniques learned in a given situation in similar or different, untrained circumstances. A generalization gradient is commonly used to express the level of generalization. Therefore, we decided to content-analyse the students' answers with respect to the following levels: 1) generalization to untrained situations; 2) generalization to similar situations; 3) no generalization/no answer.

Questionnaire Q2 includes questions related to both situations similar to the ones dealt with during the TLSs phases, and situations that appear to an expert analogous to the TLSs ones but that may not be perceived by students as similar to the TLSs' ones. We report here also the results obtained from the analysis of the pre-instruction administration of the questionnaire, even if the information that can be obtained from that data does not regard the generalization skills due to the TLSs development. On the other hand, we acknowledge that all the students have already fronted and studied some aspects of surface phenomena, during the chemistry lessons, at school. For that reason,

TABLE XIII. – $\chi^2 = 216.3$, $p < 1\%$.

Group A – Q2	Untrained situations	Similar situations	No generalization/ No answer
Pre-instruction	0	30	190
Post-instruction	40	90	39
Post-Post-instruction	31	97	41

TABLE XIV. – $\chi^2 = 230.7$, $p < 1\%$.

Group A – Q2	Untrained situations	Similar situations	No generalization/ No answer
Pre-instruction	0	31	158
Post-instruction	64	93	29
Post-Post-instruction	54	96	41

we wanted to see what was their initial generalization level with respect to situations that they could have dealt with before.

Tables XIII and XIV show the results of the analysis for Group A and Group B, respectively, in the three administration phases of the questionnaire. The values represent the number of answers that highlight a given generalization level.

As we see in fig. 6, in the pre-instruction questionnaire, the percentage of generalization in untrained situations is equal to zero, and the percentage of no generalization/no answer is significantly high for both groups. Before instruction, only a small percentage of students seem capable of generalization in situations similar to those with which they are familiar with. In answering the post-instruction questionnaire, an increase in the percentage of students who generalize in both untrained situations and situations similar to those with which they are familiar with is registered. In particular, Group B students seem to be oriented towards generalization in untrained situations more than Group A students. It is worth noting that also after instruction, a percentage not negligible of students is still unable to generalize. This is a sign that probably more time should be devoted in the TLS to modelling activities, discussion and procedures that foster extension of the learning concepts and skills to different context.

3·4. *Thematic analysis of students' contributions during the final day brainstorming phase (all variables).* – The example of results of data analysis reported here consists of the thematic analysis of the audio recordings of students' contributions during the final day brainstorming phase (database 3). In this phase, researchers suggested some topics and aspects of experimentation activities to think about and let the two groups discuss and compare their experiences.

Figure 7 shows the percentages of recurrence of each variable during the final day brainstorming phase for Group A and Group B, respectively, found according to the methods discussed in sect. 2.5.2. All variables (except variable 1.3) are studied in this analysis, but we will briefly discuss here mainly the variables related to the third dimension of learning we study in this research: Development of a mindset suited to learning science (see fig. 1).

The results of the analysis of this database show that Group B students, after trialling, have developed metacognition (3.3), well-being in learning (3.4), and willingness to extend studies and research (3.6) more than Group A students.

It is also worth noting that, after trialling, Group A students seem to have achieved appropriation of concepts and forms of representation (1.1) higher than Group B students. Moreover, at the end of the educational path, Group A show an enhancement of interpersonal and social skills (2.1) and the ability to recognize and recognize the evolution of personal cognitive styles (2.4) higher than Group B.

4. – Conclusions

In this paper, we provided an illustration of some research methods utilized by the University of Palermo Physics Education Research Group (UoP-PERG) and of the results that can be obtained by using them to study students' learning. We used the process of

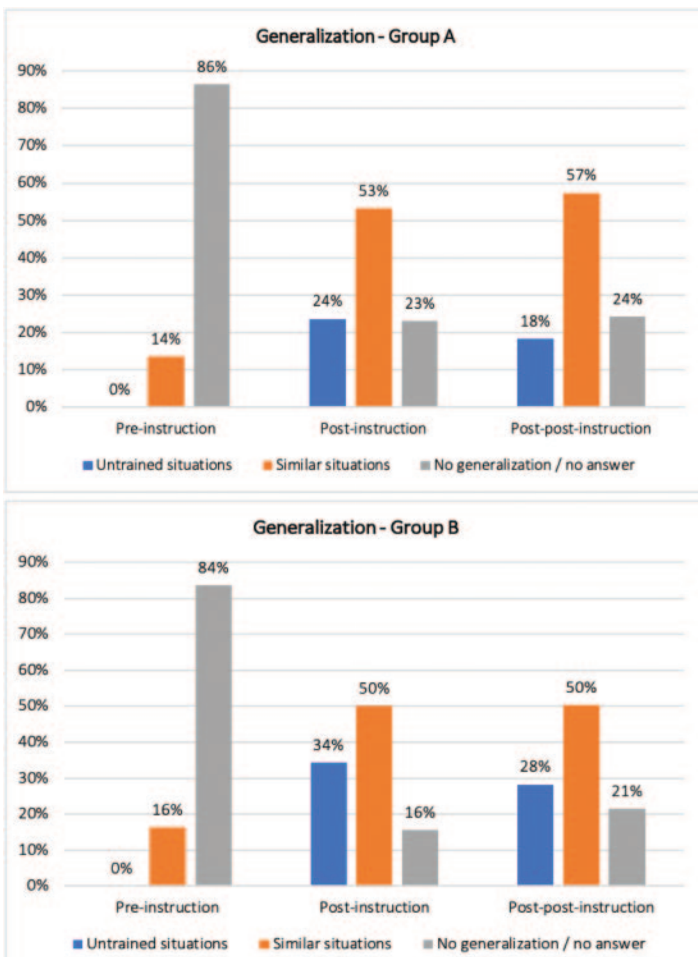

Fig. 6. – Bar diagrams showing the percentages of answers given by Group A and Group B students, respectively, to questionnaire Q2 during the pre-, post-, and post-post-instruction administrations. Answers are categorized according to students' different levels of generalization.

designing and developing two TLSs on surface phenomena in liquids as an example, to be trialled with groups of high school students. The data collection and analysis methods used in trialling the TLSs were also described.

In the introduction, we briefly discussed the educational reconstruction model, which we employed as a theoretical framework for the design of the TLSs, and a pedagogical methodology that we followed, *i.e.*, active learning. Some relevant aspects of active learning pedagogy and cognitive psychology were discussed and their relevance in the TLSs planning and development was pointed out.

As the main aim of the research was to study what aspects of the TLSs could be considered relevant in fostering significant learning in students, we provided a research-based conceptual scheme of what we mean by "improvement of students' learning", that allows us to define specific study variables. Then, we presented the TLSs phases, which are based on the well-known active-learning, inquiry- and investigation-based (ISLE) approaches. We described some methods we employed in our research to gather and analyze the data required to study the progression of student learning, with respect to

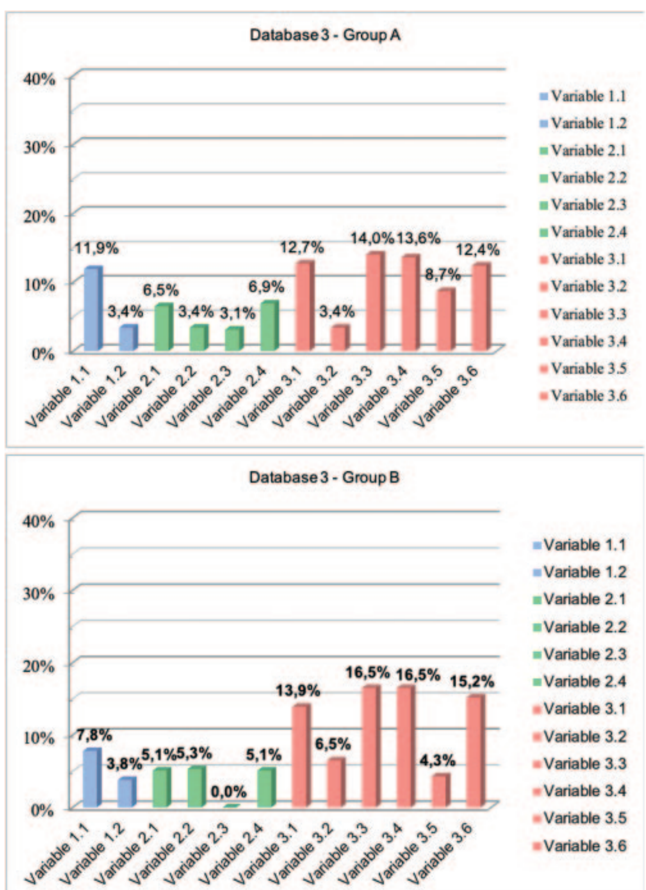

Fig. 7. – Bar-diagrams showing the percentages of recurrence of each variable during the final day brainstorming phase for Group A and Group B, respectively.

the conceptual scheme of learning provided and some specific study variables. Finally, we gave an example of the results that can be obtained with respect to some of the study variables.

The analysis of all the data collected during trialling is still in progress. Some preliminary results have already been published [92], and a complete analysis of them is in preparation. From those data, it seems that to deal with surface phenomena effectively, it should be useful to build a TLS embedding aspects of both the approaches trialled. The strong experimental connotation common to both TLSs, which in several cases required the active involvement of the students in designing experiments and collecting and analyzing data, the importance given to the continuous interaction of the students within the small and large groups, and the constant request to the students to express their agreement with the conclusions reached by the group, represented a characterizing aspect of the approach followed during the development of the two TLSs. All the students showed appreciable improvements in relation to many of the study variables. On the other hand, the different modelling approaches and the related differences in student learning showed that mesoscopic modelling seems to provide students with more effective tools for developing explanatory reasoning models with respect to the more traditional macroscopic modelling.

We also believe that presenting a topic through a multiple perspective, involving students in a complex and variegated learning environment can really help them achieve a deep understanding and awareness of themselves. All this also fosters students' acquisition of ways of reasoning and attitude towards problem-solving that represent transversal tools crucial in everyday life beyond the didactic experience. In particular, we believe that in a future trialling of this TLS it will be worth dedicating more time to the mesoscopic approach as this, despite the limited time available during the first trialling, clearly shown particular effectiveness in promoting students' understanding of the functioning mechanisms underlying the physical phenomena explained, which is essential for the achievement of scientific knowledge.

The study presented in this paper is affected by some limitations related to the number and nature of the research sample. Since the TLSs designed were tested on a sample of 40 students, all of them attending the upper secondary school, the results of our study cannot be easily generalized.

Another aspect that is worth of consideration regards the issues related to the preparation of physics teachers to design, implement and evaluate TLSs on surface phenomena, like the ones we described in this research, and, more generally, the preparation of teachers in effectively designing, implementing and evaluating pedagogical approaches based on active learning methodologies and investigation/inquiry-based approaches. The two teachers that participated in our research are both graduated in physics and interested in innovation in teaching and learning the subject. They were able to not only actively support the researchers in the design and implementation phases of the TLSs, but were also willing to transfer the pedagogical methods used during trialling to their daily teaching practice. This undoubtedly facilitated the development of the research, as we could count on the fact that the activities carried out in the classroom were systematically resumed and extended during the normal classroom activities. Unfortunately, this is not a common situation in Italian schools, where many teachers are not graduated in physics and/or do not always consider the use of active learning methodologies as feasible and useful, mainly because of time constraints with the school programs and the commitment that is required to effectively involve students in laboratory and modelling activities based on investigation/inquiry.

We plan to design a single TLS embedding the most effective aspects and tools characteristics of the macroscopic and mesoscopic approaches on surface phenomena experimented in the context of the two TLSs described in this work. On the basis of the results obtained and of the "lesson learnt", we plan to trial this upgraded TLS on a more extended sample composed of upper secondary school students (once again). A further trial of the upgraded TLS is also planned with undergraduate students (first year of Engineering degree courses) in order to get information on the influence of age and content understanding background on the improvements due to the investigative/modelling approach that can be observed with respect to the study variables.

REFERENCES

[1] Lijnse P. L., *Sci. Educ.*, **79** (1995) 189.
[2] Kattmann U., Duit R., Gropengießer H. and Komorek M., *A Model of Educational Reconstruction* (San Francisco) 1995, paper presented at The NARST annual meeting.
[3] Duit R., Gropengießer H. and Kattmann U., *Towards science education that is relevant for improving practice: The model of educational reconstruction*, in *Developing Standards in Research on Science Education*, edited by Fischer H. E. (Taylor & Francis, London) 2005, pp. 1–9.
[4] Duit R., Gropengießer H., Kattmann U., Komorek M. and Parchmann I., *The model of educational reconstruction – a framework for improving teaching and learning science*, in *Science Education Research and Practice in Europe. Cultural Perspectives in Science Education*, Vol. **5**, edited by Jorde D. and Dillon J. (Sense Publisher, Rotterdam) 2012.
[5] Psillos D. and Kariotoglou P. (Editors), *Iterative Design of Teaching-Learning Sequences* (Springer Science+Business Media, Dordrech) 2016.
[6] Meheut M. and Psillos D. (Editors), *Int. J. Sci. Educ.*, **26** (2004) 515.
[7] Schmidt W. H., Wang H. C. and mcKnight C. C., *J. Curric. Stud.*, **37** (2005) 525.
[8] Jackson J., Adams R., Schwantner U. and Waters C., *The role of learning progressions in global scales* (2018), retrieved April 29, 2023 from `https://www.brookings.edu/blog/education-plus-development/2018/05/29/the-role-of-learning-progressions-in-global-scales/`.
[9] Kim H. and Care E., *Learning progressions: pathways for 21st century teaching and learning* (2018), retrieved April 29, 2023 from `https://www.brookings.edu/blog/education-plus-development/2018/03/27/learning-progressions-pathways-for-21st-century-teaching-and-learning/`.
[10] Yew E. H. J. and Goh K., *Health Prof. Educ.*, **2** (2016) 75.
[11] Duran M. and Dokme I., *Eurasia J. Math. Sci. Technol. Educ.*, **12** (2016) 2887.
[12] Oja K. J., *J. Nurs. Educ.*, **50** (2011) 145.
[13] Herrington J. and Oliver R., *Educ. Technol. Res. Dev.*, **48** (2000) 23.
[14] Etkina E., Brookes D. T. and Planinšič G., *Investigative Science Learning Environment. When Learning Physics Mirrors Doing Physics* (IOP Publishing, Bristol, UK) 2019.
[15] Dweck C., *Mindset: The New Psychology of Success* (Ballantine Books) 2006.
[16] Ericsson K. A., *Int. J. Sport Psychol.*, **38** (2007) 4.
[17] Scherr R., *Am. J. Phys.*, **75** (2007) 272.
[18] Canning E. A., Muenks K., Green D. J. and Murphy M. C., *Sci. Adv.*, **5** (2019) eaau4734.
[19] Muenks K., Canning E. A., LaCosse J., Green D. J., Zirkel S., Garcia J. A. and Murphy M. C., *J. Exp. Psychol.: Gen.*, **149** (2020) 2119.
[20] Yeager D. S., Hanselman P., Walton G. M. *et al.*, *Nature*, **573** (2019) 364.
[21] Yeager D. S., Carrol J. M., Buontempo J., Cimpian A., Woody S., Crosnow R. *et al.*, *Psychol. Sci.*, **33** (2021) 18.

[22] Mayer R. E., *Learning and Instruction* (Pearson Education, Inc, Upper Saddle River) 2008.
[23] Barbe W. B., Swassing R. H. and Milone M. N., *Teaching through Modality Strengths: Concepts Practices* (Zaner-Bloser, Columbus) 1979.
[24] Fleming N. D. and Mills C., *Improv. Acad.*, **11** (1992) 137.
[25] Fleming N. D., *I'm different; not dumb. Modes of presentation (VARK) in the tertiary classroom*, in *Research and Development in Higher Education, Proceedings of the 1995 Annual Conference of the Higher Education and Research Development Society of Australasia*, Vol. **18** (HERDSA) 1995, pp. 308–313.
[26] Felder R. M. and Silverman L. K., *Eng. Educ.*, **78** (1988) 674.
[27] Marchand A., Weijs J. H., Snoeijer J. H. and Andreotti B., *Am. J. Phys.*, **79** (2011) 999.
[28] Develaki M., *Sci. Educ.*, **26** (2017) 1001.
[29] Besson U. and Viennot L., *Int. J. Sci. Educ.*, **26** (2004) 1083.
[30] Besson U., *Sci. Educ.*, **19** (2010) 225.
[31] Marzano R. J., *A Different Kind of Classroom: Teaching with Dimensions of Learning* (Hawker Brownlow Education) 2006.
[32] Levrini O., Fantini P., Tasquier G., Pecori B. and Levin M., *J. Learn. Sci.*, **24** (2015) 93.
[33] Rogoff B., *Observing sociocultural activity on three plans: Participatory appropriation, guided participation and apprenticeship*, in *Sociocultural Studies of Mind*, edited by Wertsch J. V., del Rio P. and Alvare A. (Cambridge University Press, New York, NY) 1995, pp. 139–164.
[34] DiSessa A. A. and Sherin B. L., *Int. J. Sci. Educ.*, **20** (1998) 1155.
[35] Ogborn J., *Rev. Bras. Pesqui. Educ. Ciênc.*, **6** (2011) 1.
[36] Magnusson S. J., Templin M. and Boyle R. A., *J. Learn. Sci.*, **6** (1997) 91.
[37] Kastrinos W., *Sci. Educ.*, **49** (1965) 487.
[38] Arthur W. jr., Bennett W. jr., Stanush P. L. and McNelly T. L., *Hum. Perform.*, **11** (1998) 57.
[39] Streveler R. A., Litzinger T. A., Miller R. L. and Steif P. S., *J. Eng. Educ.*, **97** (2008) 279.
[40] Gresham F. M., *J. Psychoeduc. Assess.*, **1** (1983) 299.
[41] McFall R. M., *Behav. Assess.*, **4** (1982) 1.
[42] Fazio C., Battaglia O. R., and Di Paola B., *Phys. Rev. ST-Phys. Educ. Res.*, **9** (2013) 020101.
[43] Bao L., and Redish E. F., *Phys. Rev. ST-Phys. Educ. Res.*, **2** (2006) 010103.
[44] Gluck M. A., Mercado E. and Myers C. E., *Learning and Memory: From Brain to Behavior* (Worth Publishers, New York) 2008, p. 640.
[45] Banich M. T. and Caccamise D., *Generalization of Knowledge: Multidisciplinary Perspectives*, 1st edition (Psychology Press) 2010.
[46] Shea T. M. and Bauer A. M., *Behavior Management: A Practical Approach for Educators* (Pearson Education) 2007.
[47] Saddler-Smith E., *Pers. Indiv. Diff.*, **30** (2001) 609.
[48] Bandura A., *Educ. Psychol.*, **28** (1993) 117.
[49] Bandura A., *Self-Efficacy: The Exercise of Control* (W. H. Freeman, New York, NY, USA) 1997.
[50] Berry J. M., *A Self-Efficacy Model of Memory Performance* (American Psychological Association, New York) 1987.
[51] Zimmerman B. J., Bandura A. and Martinez-Pons M., *Am. Educ. Res. J.*, **29** (1992) 663.
[52] Zhang J., Kuusisto E. and Tirri K., *Psychology*, **8** (2017) 1363.
[53] Baker L., *Metacognition in comprehension instruction*, in *Comprehension Instruction: Research-Based Best Practices* (Guilford Press, New York) 2002, pp. 77–95.
[54] Harris K. R., Graham S., Brindle M. and Sandmel K., *Handb. Metacogn. Educ.*, **131** (2009) 153.

[55] Noddings N., *Happiness and Education* (Cambridge University Press) 2003.
[56] Konu A. and Rimpelä M., *Health Promot. Int.*, **17** (2002) 79.
[57] Sellström E. and Bremberg S., *J. Epidemiol. Community Health*, **60** (2006) 149.
[58] Boekaerts M., *Educ. Phychol.*, **28** (1993) 149.
[59] Schwartz R. S., Lederman N. G. and Crawford B. A., *Sci. Educ.*, **88** (2004) 610.
[60] Abd-El-Khalick F., *Int. J. Sci. Educ.*, **34** (2012) 353.
[61] McDevitt B., *Lang. Learn. J.*, **16** (1997) 34.
[62] Freeman S., Eddy S. L., McDonough M., Smith M. K., Okoroafor N., Jordt H. and Wenderoth M. P., *Proc. Natl. Acad. Sci. U.S.A.*, **111** (2014) 8410.
[63] Jensen M., *Clin. J. Pain*, **19** (2003) 345.
[64] Wenning C. J., *J. Phys. Techr. Educ. Online*, **2** (2005) 3.
[65] McDermott L. C., Shaffer P. S. and Constantinou C. P., *Phys. Educ.*, **35** (2000) 411.
[66] Battaglia O. R., Agliolo Gallitto A., Termini G. and Fazio C., *Eur. J. Phys.*, **44** (2023) 055001.
[67] Specht M., Frömbgen T., Wessely H. and Sokolowski M., *Am. J. Phys.*, **87** (2019) 1014.
[68] `https://www.youtube.com/watch?v=IjD5e33FnP0`.
[69] `https://www.youtube.com/watch?v=iBvKJyT5Jww`.
[70] `https://www.youtube.com/watch?v=tZrI-5gTv9o`.
[71] `https://www.youtube.com/watch?v=7Who8EpbvCY`.
[72] `https://www.youtube.com/watch?v=pmagWO-kQOM`.
[73] `https://www.youtube.com/watch?v=5NCOnr3VSAY`.
[74] Battaglia O. R., Agliolo Gallitto A. and Fazio C., *Comput. Appl. Eng. Educ.*, **27** (2019) 1333.
[75] Battaglia O. R., Agliolo Gallitto A., Termini G. and Fazio C., *Eur. J. Phys.*, **42** (2021) 055704.
[76] Monaghan J. J., *Rep. Prog. Phys.*, **68** (2005) 1703.
[77] Marton F., *J. Thought*, **21** (1986) 28.
[78] Krippendorff K., *Content Analysis: An Introduction to its Methodology* (Sage Publications) 2018.
[79] Battaglia O. R., Di Paola B. and Fazio C., *Phys. Rev. Phys. Educ. Res.*, **15** (2019) 020112.
[80] Fazio C., Battaglia O. R. and Di Paola B., *Phys. Rev. ST-Phys. Educ. Res.*, **9** (2013) 020101.
[81] Boyatzis R. E., *Transforming Qualitative Information: Thematic Analysis and Code Development* (Sage) 1998.
[82] Braun V. and Clarke V., *Qual. Res. Psychol.*, **3** (2006) 77.
[83] Battaglia O. R., Di Paola B., Persano Adorno D., Pizzolato N. and Fazio C., *Res. Sci. Educ.*, **49** (2019) 1395.
[84] Everitt B. S., Landau S., Leese M. and Stahl D., *Cluster Analysis* (John Wiley and Sons, Ltd, Chichester, UK) 2011.
[85] Ott J., *Analysis of Human Genetic Linkage*, 3rd edition (Johns Hopkins University Press, Baltimore, London) 1999.
[86] Allen D. N. and Goldstein G. (Editors), *Cluster Analysis in Neuropsychological Research: 13 Recent Applications* (Springer Science+Business Media, New York) 2013.
[87] Mantegna R. N., *Eur. Phys. J. B*, **11** (1999) 193.
[88] Cowgill M. C. and Harvey R. J., *Comput. Math. Appl.*, **37** (1999) 99.
[89] Dayan P., *Unsupervised Learning*, in *The MIT Encyclopedia of the Cognitive Sciences*, edited by Wilson R. A. and Keil F. (The MIT Press) 1999, pp. 1–7.
[90] Sathya R. and Abraham A., *Int. J. Adv. Res. Artif. Intell.*, **2** (2013) 34.
[91] Battaglia O. R., Di Paola B. and Fazio C., *Phys. Rev. Phys. Educ. Res.*, **15** (2019) 020112.
[92] Battaglia O. R., Agliolo Gallitto A., Termini G. and Fazio C., *Educ. Sci.*, **13** (2023) 425.