

Long term stability of learning outcomes in undergraduates after an open-inquiry instruction on thermal science

Dominique Persano Adorno, Nicola Pizzolato,^{*} and Claudio Fazio
*UOP-PERG (University of Palermo-Physics Education Research Group),
 Dipartimento di Fisica e Chimica, Università di Palermo, 90128 Palermo, Italy*

 (Received 30 April 2017; revised manuscript received 9 October 2017; published 7 February 2018)

This paper investigates the efficacy of an open-inquiry approach to achieve a long term stability of physics instruction. This study represents the natural continuation of a research project started four years ago when a sample of thirty engineering undergraduates, having already attended traditional university physics instruction, were involved in a six-week long learning experience of open-inquiry research activities within the highly motivating context of developing a thermodynamically efficient space base on Mars. Four years later, we explore the effectiveness of that learning experience by analyzing the outcomes that the students achieved by answering again the same questionnaire that was administered them both prior to and immediately after those activities. As we did in the first work, students' answers were classified within three epistemological profiles. Now, a comparison among students' outcomes during the three phases, namely, preinstruction, postinstruction, and after four years has been carried out. Immediately after the open-inquiry experience, the students obtained significant benefits in terms of the strengthening of their practical and reasoning abilities, by proficiently applying the learned concepts to face and solve real-world problem situations. In this study, the students' answers do not highlight any significant regress towards their preinstruction profiles. The global robustness of the teaching strategy adopted four years ago is confirmed by a statistically significant comparison with a control group of students who experienced the same curricular instruction except for the open inquiry-based workshop. Nevertheless, some changes have been observed and discussed in the light of the answers the students provided to a short interview regarding their studying or working experiences across the four-year temporal window.

DOI: [10.1103/PhysRevPhysEducRes.14.010108](https://doi.org/10.1103/PhysRevPhysEducRes.14.010108)

I. INTRODUCTION

The achievement of long lasting and effective learning is the main purpose of any scientific instruction [1–3]. This is important for everyone in order to approach everyday life in a more conscious way, and particularly for all those students directly engaged in science or engineering practices. For them, an effective science instruction should support the development of useable knowledge in order to foster the application of core ideas across the science disciplines [4–6]. Moreover, through developing and using models, planning and conducting investigations, analyzing and interpreting data, using mathematical and computational thinking, and constructing explanations, future scientists should gain a deeper understanding of fundamental disciplinary concepts and, at the same time, strengthen

reasoning skills and transversal abilities in solving everyday life problems [7–9]. For these reasons, a scientific instruction should be considered most effective when it produces a lifelong lasting learning [10,11].

In this view, science education carried out by fostering the development of the process of *inquiry*—learning through questioning—and the practice of scientific reasoning is nowadays considered the most effective framework for teaching or learning science, in terms of an active construction of long term meaningful knowledge, supporting the overcoming of both conceptual and epistemological difficulties in problem solving [12–15]. An inquiry-based science education is becoming a common teaching paradigm all over the world at both primary and secondary school. Unfortunately, except for some notable experiences of inquiry-based instruction, such as the Workshop Physics [16,17], Real-Time Physics [18–20], Interactive Lecture Demonstrations [19], and Conceptual Labs [21], this is generally hardly happening at university, where most scientific disciplines are still taught by following a lecture-based method of instruction. In some cases, this might be due to greater numbers of students attending a university course, limiting the possibility to engage them within an inquiry-based learning environment in real laboratories.

^{*} nicola.pizzolato@unipa.it

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

TABLE I. Summary of characteristics for the three student answering typologies reported in Ref. [23].

Practical or everyday (pe-type)	Descriptive (de-type)	Explicative (ex-type)
The student faces practical physics problems by using only their commonsense experience, providing examples from everyday life and, in some cases, by mentioning physical quantities which are not useful to achieve a coherent explanation of the described phenomenon.	The student approaches the problem resolution by focusing on their background of knowledge, by writing one or more mathematical expressions in order to achieve a reasonable description of the observed phenomenon, but without any further step towards the problem resolution.	The student achieves a successful explanation of the proposed problem by effectively applying their theoretical background of knowledge within the appropriate physical context, addressing the most relevant physical quantity involved in the problem and providing a reasonable explicative model.

In other cases, laboratories are still considered places to show demonstrative experiments, having previously introduced the students to a physics theory or law [22].

In a previous study we addressed the problem of developing an effective strategy of instruction for science or engineering students by involving a group of engineering undergraduates, who had already attended regular physics lectures, to experience an open-inquiry (OI) based learning environment on thermal phenomena [23]. A questionnaire, specifically developed to investigate the presence of problem-solving related epistemological difficulties affecting the deep understanding of common life phenomena in thermal science, was administered to the students both prior to and after the OI experience. Even if students had already attended traditional courses on thermal science before the OI-based workshop, a significant fraction of students experienced difficulties in applying physics concepts to solve problems. In the framework provided by domain-specific expertise and by following a phenomenographic approach [24–27] to the analysis of students' answers to the questionnaire, it was possible to categorize typical answering strategies (see Table I in Ref. [23] and here summarized in Table I), identifying three ideal student profiles: (i) the practical or everyday (pe)-type, characterizing the student who remains embedded in the everyday common sense explanation without achieving a scientific comprehension of the observed phenomena; (ii) the descriptive (de)-type, associated with the learner who studies physics laws and theories in a purely mnemonic formulation; and (iii) the explicative (ex)-type, referring to the ability of finding explicative models and effective reasoning for the resolution of common life scientific problems.

The statistical analysis of the similarity relationships between the real students' answering strategies and the three ideal student profiles allowed us to quantify the positive effects of the OI-based learning experience on the student epistemological stances with respect to facing thermal phenomena and solving related problems [23]. As a matter of fact, at the beginning of the workshop the students showed higher levels of similarity with the de-type profile, while at the end of OI-based activities the majority of them adopted ex-type reasoning strategies.

In this work we report the results of a study aimed at investigating the long term effectiveness of those students' outcomes about four years after they attended the OI-based workshop. Our previous work was significant in that it was one of the first to explore quantitative relationships between inquiry-based instruction and student epistemologies. Here, for the first time to the best of our knowledge, a quantitative analysis about the long term stability of student strengthening of reasoning skills for everyday problem solving in thermal science is presented. Only two previous studies attempted to address this issue some time ago by investigating the long term learning effects of “active engagement” methods in mechanics courses at university, providing promising results about the achievement of permanent change in students' conceptual framework up to three years after instruction [28,29]. We invited the same students from the first study—now mostly graduated professionals—to answer the same questionnaire they answered four years ago in the research-like context of the “Mission to Mars” OI-based workshop. A comparative study between the answering strategies implemented by the students now and those provided four years ago has been carried out by applying the same method of similarity analysis performed in the previous study, with respect to the three ideal student profiles described above. A control group of students, selected for having experienced the same curricular path of instruction with the research sample, except for the OI-based experience, has been used for comparison in order to strengthen our findings. In the following section we introduce a short background, the rationale of our research study, and the research question addressed in this paper. The method of sample selection, data collection and analysis, and the results and discussion are reported in the subsequent sections. The conclusion section reports a summary of the results and final comments about the advantages and limits of OI-based teaching strategies for a long lasting physics education of science and engineering students.

II. THEORETICAL BACKGROUND AND RESEARCH RATIONALE

Teaching science, and physics in particular, has always had a twofold objective. First, it should provide every

person with an adequate wealth of knowledge, in terms of both interdisciplinary scientific concepts and practical abilities in order to recognize and fix scientific troubles they might face in the context of everyday life. Second, in particular for all those students who want to undertake a career in science and consider the opportunity to make scientific research their future job, the focus of an effective teaching process should be the strengthening of the reasoning skills of the student, who needs to be suitably trained to develop a scientific mind. Solving problems by using scientific concepts effectively has always been considered a critical part of this scientific upbringing. In this regard, a fundamental role is played by the inquiry process, i.e., the reasoned questioning that drives any sort of scientific research. In inquiry-based instruction, the students are engaged in highly motivating learning environments and focused on a specific problem to be addressed by following a circle of activities [30,31]: identify scientifically oriented questions, plan investigations, collect evidences, build explanation models, share findings, and eventually address new questions that might arise. In this context, depending on the amount of information and support provided by the teachers, the learners may be involved in a structured or guided inquiry or in OI [32]. In OI-based instruction, the teacher takes the delicate role of defining the context for inquiry, stimulating the students to derive their own questions about a problem, design and carry out independent investigations, construct coherent explanations, and discuss their findings. This level of inquiry requires higher capacities of scientific reasoning with respect to a more structured and guided inquiry, but, for this reason, it is considered the best training to achieve higher levels of critical thinking skills, involving flexibility, judgment, and contemplation, as part of the changes that occur in the course of inquiry [33–36]. However, the introduction of the students to an OI learning environment without any prior instruction on a specific topic might require an excess of cognitive effort from some of them in terms of scientific reasoning, inducing feelings of inadequacy or frustration, and could not bring about an effective understanding of the concepts [37–39]. Recently, an integrated teaching method with students approaching an OI-based instruction after a traditional lecture-based course seems to offer a valuable possibility of effective teaching and learning strategy [23].

Students involved in OI-based learning experiences have the authentic opportunity to perform researchlike activities as real scientists. Definitely, this contributes to achieving a robust awareness of the process of scientific inquiry and a deeper view of the nature of science [40–43]. As a matter of fact, engineering students, in particular, are too often focused on mechanically solving standard problems with the only aim of passing the exam, encountering many difficulties when facing nonstandard problems about everyday phenomena. They have no idea about the way scientific

issues are addressed by real scientists and how complex phenomena can be investigated. Reasoning efforts are needed to overcome the barriers of the everyday thinking, simplify the analysis of a complex event by means of simplified models, apply conceptual knowledge effectively, and provide meaningful explanations. The understanding on how scientific knowledge is produced in real research contexts, jointly with the way inquiry is implemented in the classroom, affects the epistemological ideas that students might have about their learning activity, changing their perspective on scientific knowledge and favoring their ability to step forward to finding a solution to a problem [40–42].

Student epistemologies of science are not robust beliefs that drives students' learning and problem solving, but context-dependent locally coherent views whose stability depends both on external stimuli and on students' personal conceptions and emotional states [44,45]. In this context, the lasting of such changes over a long temporal step after an inquiry-based instruction has never been explored before.

The research question used to guide this work is

“How effective is an OI-based instruction to achieve a long term stability of undergraduates' reasoning skills and changes in student epistemologies to solve problems concerning everyday thermal phenomena?”

The management of many science or engineering problems requires high-level thinking skills. Four years ago the topic of thermal physics was selected among those particularly requiring a meaningful understanding of complex physics concepts typical of this branch of science [46–50]. However, any instruction cannot be considered fully effective until a long term stability of students' outcomes is achieved, considering the necessities to apply the gained reasoning abilities to the resolution of everyday problems and future research issues.

III. METHOD

In the previous research, thirty engineering undergraduates, after having attended regular lectures on thermal science, were involved in a six-week long OI-based learning experience; specifically, a highly motivating simulated research context regarding the design of a thermodynamically efficient space base on Mars [23]. Those students had never been involved in inquiry-based learning paths before that time. The processes of scientific thinking connected with their epistemological beliefs on how to address science questions were investigated by using a questionnaire with common life open-ended problems on thermal science, which was administered both prior to and after the OI-based instruction. The robustness of the questionnaire in terms of internal validity and reliability was confirmed in the first study. The students were not

TABLE II. Number of students having specific characteristics of instruction in this study.

GROUP 1 First level degree (3 years)	GROUP 2 Master degree student	GROUP 3 Master degree (3 + 2 years)	GROUP 4 Ph.D. Student or research fellow
7 (23.3%)	6 (20%)	6 (20%)	11 (36.7%)
Student label 2, 5, 7, 21, 23, 26, 29	Student label 9, 10, 14, 16, 18, 25	Student label 4, 6, 12, 17, 24, 27	Student label 1, 3, 8, 11, 13, 15, 19, 20, 22, 28, 30

informed about the duty to answer to the same questionnaire both prior to and after instruction (and four years later). None of the problems proposed in the questionnaire items were explicitly addressed during the OI-based learning workshop. For this reason, we considered the modifications in the students' postinstruction answers as mainly due to their personal revisions of the involved situations, as a consequence of the proposed learning path.

In the present study, the same questionnaire has been submitted to the same group of students and the results compared with those collected four years before. The modality of this latter administration of the questionnaire has been the same of the previous two carried out four years ago (face-to-face administration). The administration of the questionnaire has been followed by a short structured interview where the students in person were asked to answer, in order to collect a feedback on the continuation of their instruction path after the OI experience. In Table II we report the number of students as divided into four groups characterized by different levels of instruction, specifically: First level degree (G1), grouping those students who earned the 3-year degree and quit studying, Master's degree student (G2), with students still involved into Master's degree instruction, Master's degree (G3), including graduated engineers now working, and Ph.D. student or research fellow (G4), grouping graduated professionals involved in scientific research. Being useful for the subsequent analysis, in Table II we also specified the student label as recorded in Ref. [23].

The information collected by interviewing our research sample helped us to select a control group of students who have followed the same curricular path of instruction as the intervention group, but have not participated in the open-inquiry instruction in thermal science four years ago. In particular, the control group has been arranged by selecting a total of thirty among engineering undergraduates and graduated engineers grouped within the same partition as the research sample in Table II. The selection process of the participants to the control group has been carried out by calling the suitable candidates identified on the basis of their records at university (by selecting those with marks as close as possible to the average values of marks of the students as grouped in Table II) and by asking their availability to join in this study. The participants in the control group answered the same questionnaire (with the same modalities) administered to the research sample.

The same method of statistical implicative analysis [11,51,52] used in the previous work [23] has been applied here again to quantitatively estimate the similarities between different students' answering strategies and the three ideal epistemological profiles, characterized by different levels of ability to face and solve real-world problems by proficiently applying the thermal physics laws. The analysis of the students' outcomes has been performed by means of the C. H. I. C. software [53], which has been used to quantify the similarity relationship of different problem-solving approaches to ideal epistemological profiles. The Lerman's similarity index, originally defined in Ref. [11] to classify student behavior according to a method of hierarchical clustering [51] and used in the previous paper to recognize similar answering strategies (see Sec. III. B in Ref. [23]), is also used here to quantify the similarity (in a range 0–1) between answering strategies of each student of the research sample or control group and the three ideal epistemological profiles.

IV. RESULTS AND DISCUSSION

In this section we report and discuss the results obtained by analyzing the answers provided to the questionnaire by our research sample of engineering students both prior to (preinstruction), immediately after (postinstruction), and four years later (postinstruction 4y) the experience of the OI-based learning workshop. These results are discussed in comparison with those obtained by analyzing the outcomes to the questionnaire provided by the control group.

In Fig. 1 we show the average similarity index of the research sample with respect to the three epistemological profiles for the preinstruction (red), postinstruction (blue), and post 4 years (green) questionnaire answering. For comparison, in the same plot we also show the average similarities obtained by the analysis of the answering strategies provided by the control group (purple). Vertical black lines indicate the distance of 2 standard deviations of the mean.

The average changes in student problem-solving epistemologies recorded immediately after the OI-based instruction (blue data) with respect to the preinstruction data (red) were extensively addressed in the previous paper [23] and will be no longer discussed here. In this work, we are interested to compare the outcomes collected four years later (green data) with the previous two (red and blue data).

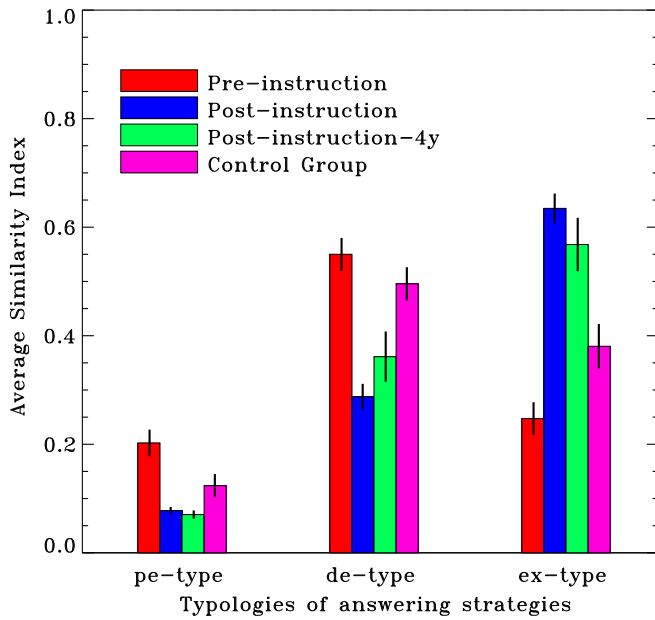


FIG. 1. Average similarity index of student answering strategies with pe-type, de-type, or ex-type profiles to address the questionnaire, administered before (red bars), just after instruction (blue bars) and four years after participating the OI-based experience (green bars), in comparison with the results obtained by the control group (purple bars).

From an overall point of view, the results obtained four years after instruction are closer to those recorded immediately after the OI-based workshop than to those obtained before instruction. The average similarity of the green data set with the pe-type is not essentially different from that of blue data set, while a slightly higher similarity index to the de-type (and a corresponding lower similarity to the ex-type) is observed.

The results obtained four years after instruction in our research sample show significant differences with the outcomes from the control group (purple data). The average similarity indexes of the purple data set to pe-type, de-type, and ex-type profiles are well outside the distance of 1 standard deviation from the mean calculated by the green data. In order to get a statistical significance of the difference between the mean of the two samples (green and purple) we have computed the student's *T* statistic separately for the three answering typologies, and the results are summarized in Table III.

TABLE III. Student's *T*-statistic results for the comparison between the research sample and control group.

Answering typologies	Practical or everyday (pe-type)	Descriptive (de-type)	Explicative (ex-type)
<i>T</i>	2.4	2.5	3.0
Significance	0.020	0.017	0.0043
Confidence level	98%	98.3%	99.5%

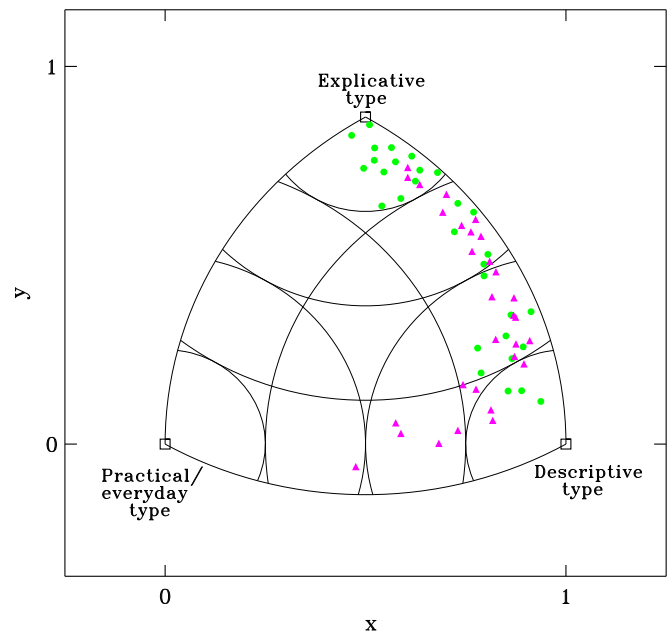


FIG. 2. Similarity graph for post-4y-instruction (green circles) and control group (purple triangles) data. The squares at the vertices of an equilateral triangle represent the pe-type, de-type, and ex-type ideal answering profiles. The *x* and *y* represent the coordinates in the similarity space. The concentric circumferences mark the distances 0.25, 0.5, 0.75, and 1 from the vertices.

Being the values of significance so low and considering the computed confidence levels are always well above the usual 95% threshold generally adopted to establish that two sample populations have significantly different means, we may infer there is a statistically significant difference between the research sample and control group—rejecting the possibility that the differences are due to random sampling variability.

The detailed distribution of student answering strategies with respect to the three ideal epistemological profiles of problem-solving approaches is shown by means of the similarity graph reported in Fig. 2, where the similarity pattern of the research sample (green circles) is compared with the control group (purple triangles), through the answers collected four years after the OI-based instruction. In Fig. 2, the three boundary squares are the vertices of an equilateral triangle with side length equal to unity and they represent the three ideal epistemological typologies of problem approaching. The distance *d* from each one of these vertices is related to the similarity index *s* by $d = 1 - s$. The *x* and *y* represent the coordinates in a similarity space where the proximity to a specific vertex implies a high value of similarity with that typology, and vice versa. The concentric circumferences mark the distances 0.25, 0.5, 0.75, and 1 from the vertices and the colored-numbered symbols represent the students.

In Fig. 2 the main difference between the two samples arises from the presence of a cluster of control group's

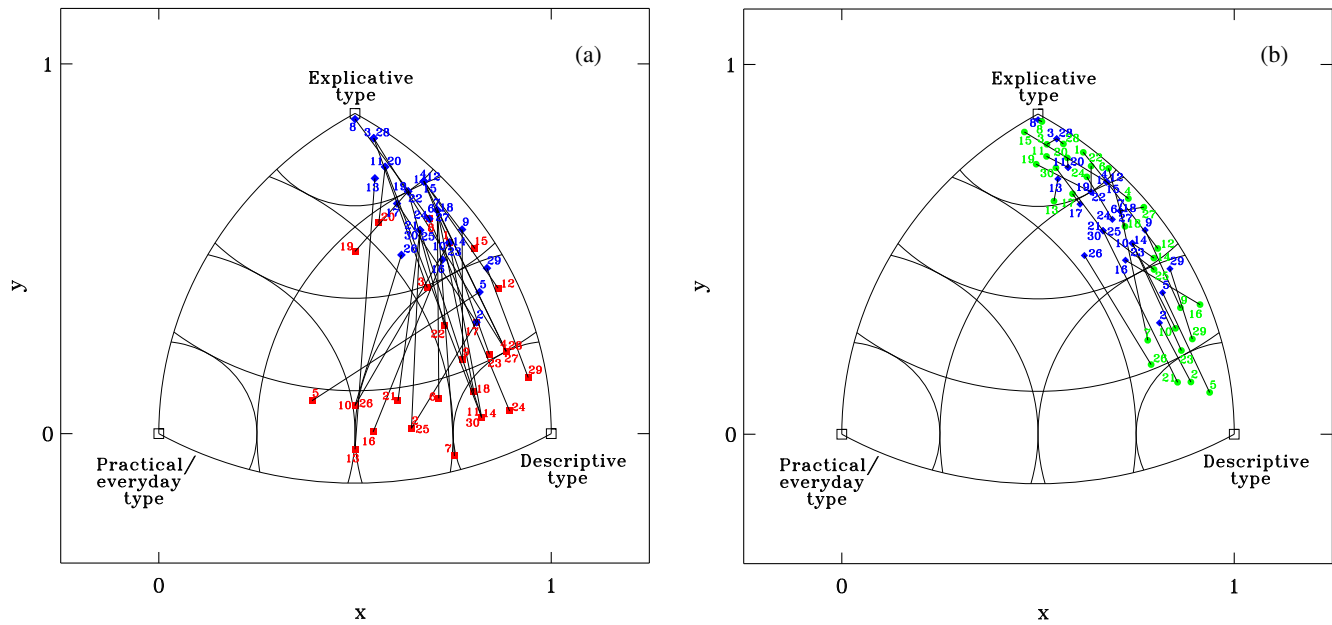


FIG. 3. Similarity graph for pre or post (a) and post or post-4y (b) instruction data. The squares at the vertices of an equilateral triangle represent the practical or everyday, descriptive and explicative typologies. The x and y represent the coordinates in the similarity space. The concentric circumferences mark the distances 0.25, 0.5, 0.75, and 1 from the vertices, the colored or numbered symbols represent the students.

participants with higher pe-type similarity index (lower distances from practical or everyday vertex) with respect to the research sample, confirming the higher mean reported in Fig. 1. Second, only three units of the control group lie within the 0.25 proximity circle to the ex-type vertex, with respect to a definitely more populated subsample of students from the intervention group. This finding makes sense of the lower average similarity index to the ex-type profile (plotted in Fig. 1) than that obtained from the research sample four years after instruction.

In Fig. 3 we show a detailed comparison between the preinstruction (red squares), postinstruction data (blue diamonds), and four-year-later outcomes (green circles). In this plot solid lines are used to connect the outcomes from any single student. In this way, we have analyzed the overall stability of the answering strategies implemented by our sample of students and, at the same time, we have followed the temporal evolution of the problem-approaching epistemology of any single participant to this study. Figure 3(a) essentially summarizes the results obtained immediately after the accomplishment of the OI-based learning experience [23]. Initially, the students' answers highlighted a combination of two main clusters of similarity, specifically around the de-type ideal profile with a less populated one related to the ex-type profile, with some students sharing similarities with both practical or everyday and descriptive typologies. This finding has been extensively discussed in our previous paper where we addressed the problem solving difficulties experienced by already-instructed undergraduates in terms of their

epistemological stances, probably ascribed to a learning method too much concentrated towards solving standard problems by applying mnemonic procedures, instead of deepening their view of science as a reasoning effort for finding coherent explanations. Postinstruction data show a strong reduction in the students' pe-type similarity index (greater distances from practical or everyday vertex) and a significant increase in the number of explicative strategies, forming a dominant ex- or de-type (blue) cluster of similarity.

The comparison between postinstruction data (blue diamonds) and four-year-later outcomes (green circles) is shown in Fig. 3(b). Immediately after the completion of the OI-based learning workshop the students answered the questionnaire by mainly applying problem solving strategies characterized by an explicative approach, as highlighted by a concentrated cluster of blue-diamond symbols, with the exception of the students labeled 2, 5, and 29, whose outcomes lie a little closer to a descriptive type. Four years later, the distribution of student answering strategies appears to be more scattered with the predominance of two distinct clusters around the explicative and descriptive typologies. In this respect, a closer look at this graph allows one to note that the change into the similarity distribution appears to follow a definite trend, with some students improving their outcomes towards the explicative corner while others seem to regress to a de-type. In order to further explore this point, we have plotted in Fig. 4 the values of similarity index for all the students of the research sample, separately for pe-type, de-type, and ex-type profiles. For each student, labeled by an integer from 1 to 30 on

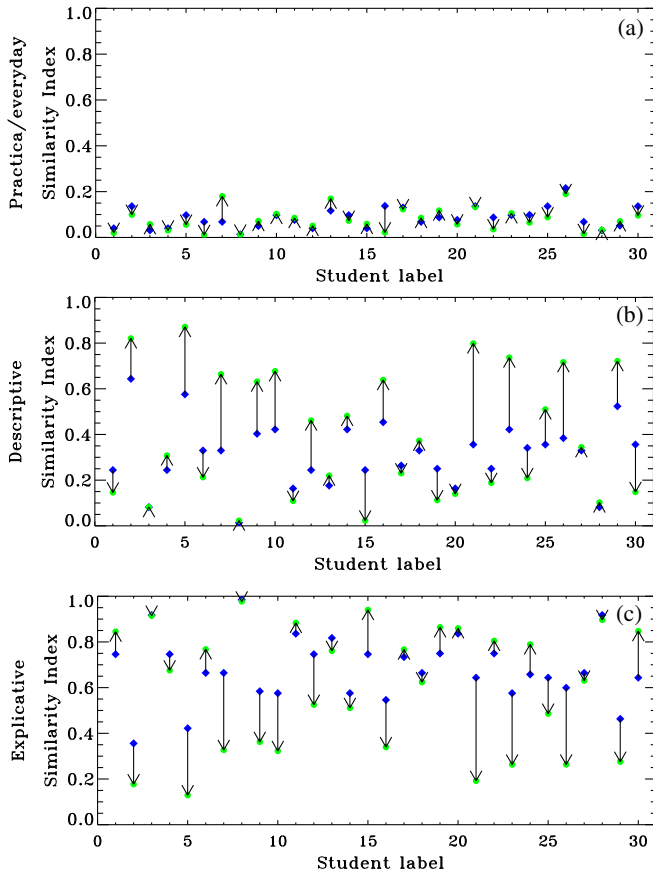


FIG. 4. Similarity indexes for all the students, each one labeled by an integer index on the abscissa, to the pe-type (a), de-type (b), and ex-type (c). For each student, blue diamonds and green circles connected by an arrow distinguish postinstruction data from those collected four years later.

the x axis, two different symbols are used to distinguish postinstruction (blue diamonds) from post-4y-instruction (green circles) similarity indexes, connected by an arrow helping to visualize the temporal evolution.

First, we have calculated the paired T statistics for the comparison between the research sample’s average similarities observed immediately after instruction and four years later (Table IV). In terms of significance, the results of this test confirm that the average similarities with respect to the practical or everyday epistemological profile did not

TABLE IV. Student paired T -statistic results for the comparison between the research sample’s outcomes collected immediately after instruction and four years later.

Answering typologies	Practical or everyday (pe-type)	Descriptive (de-type)	Explicative (ex-type)
T	1.14	2.53	2.21
Significance	0.26	0.017	0.035
Confidence level	74.0%	98.3%	96.5%

change significantly with time. As a matter of fact, data in Fig. 4(a) do not highlight significant changes in the similarities of the student answering strategies, confirming the stability of the student learning outcomes with respect to a regression towards the pe-type profile. The higher levels of confidence obtained in the two remaining profiles suggest the presence of differences which needed a detailed investigation. In fact, with the exception of some students characterized by a high level of stability (labeled 3, 8, 17, 20, 27, and 28), a greater variability is instead observed for similarities to the de-type and ex-type [Figs. 4(b) and 4(c)], where some students show an increase of similarity with the de-type approach (and a consequent decrease with respect to the ex-type) and others, on the contrary, are characterized by a reversed behavior, showing an improvement towards the ex-type answering profile. In particular, a deeper analysis of the data plotted in Figs. 4(b) and 4(c) highlights the presence of three subsamples of students whose similarity changes behaves in a similar way. A first group (with students labeled 1, 6, 11, 15, 17, 19, 22, 24, 30 in Fig. 4) shows a definite improvement in their similarity towards the explicative typology. A second group (with learners labeled 2, 5, 7, 9, 10, 14, 16, 21, 23, 25, 26, 29), characterized by an initial (postinstruction) similarity index lower than 0.6 with respect to the ex-type, has further lowered these values by increasing the similarity towards the de-type profile. A final group (with students labeled 3, 4, 8, 18, 20, 27, 28), all starting with ex-type similarities higher than 0.6, after four years, essentially remained with the same similarities with respect to the three typologies.

In order to explore the connections between the trends observed in Fig. 4 about the changes of student problem solving approaches and the instruction path the students have followed in the last four years, we have plotted in Fig. 5 a multiple similarity graph with the data collected before, immediately after the OI-based workshop, and four years after instruction, for the research sample (blue circles) and control group (red diamonds) separately as in the four groups of Table II. In this way, we can follow in Fig. 5, row by row from left to right, the temporal evolution of the problem solving approach of students within a single group and finally compare this behavior with the corresponding participants from the control group.

Even considering the small number of members in each group due to the limitation of a pilot study, some definite behavior can be easily recognized in Fig. 5. Four years after the OI-based workshop, students in G1, who did not continue their studies, show a clear change towards the de-type profile, although maintaining about the same distances from the pe-type vertex achieved immediately after the OI-based instruction. A similar but less evident behavior is recognized in the Master’s degree students belonging to G2. However, it should be noticed that, even though some regression towards the de-type is observed, the students belonging to G1 and G2 are still fairing, on

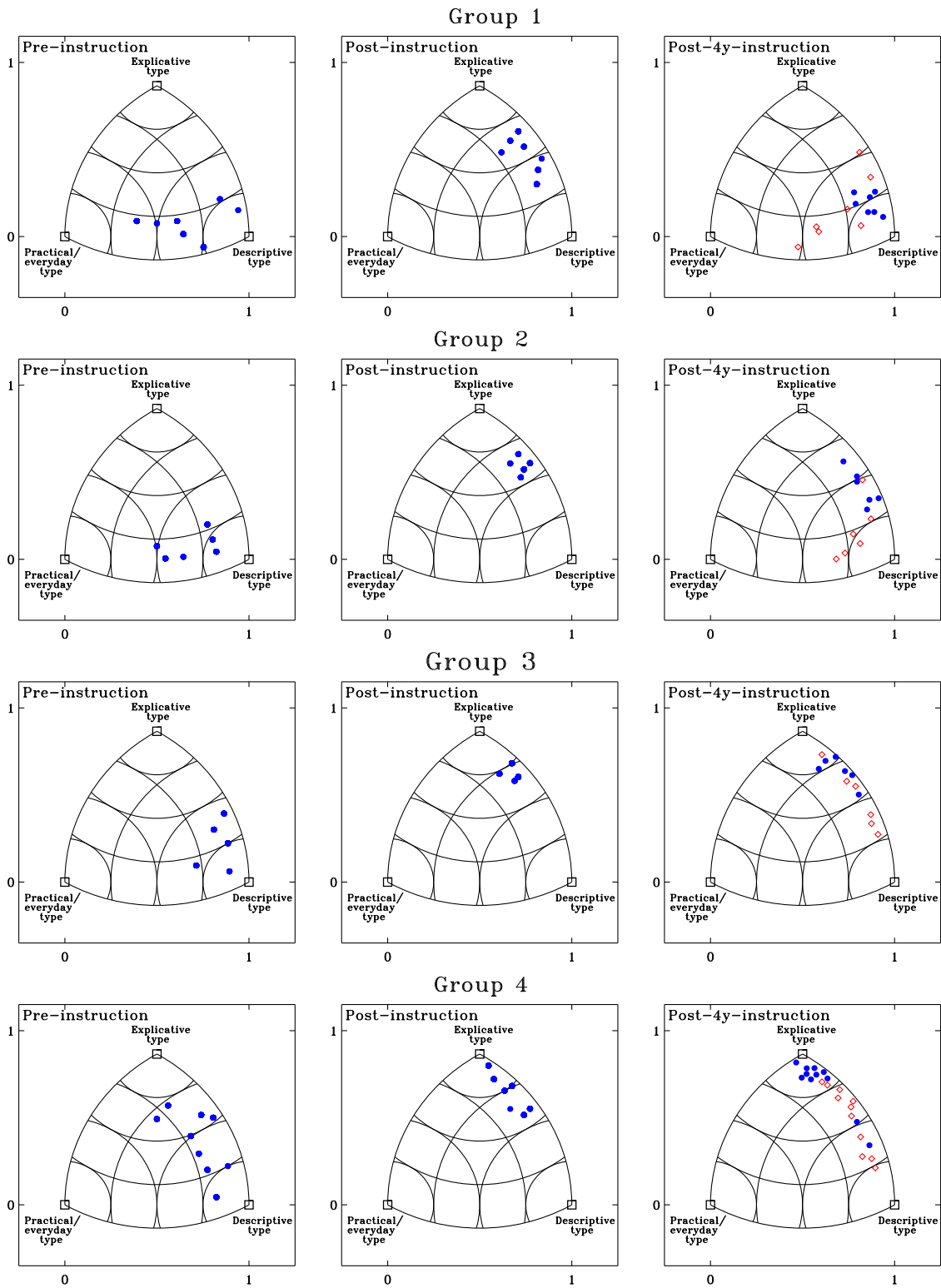


FIG. 5. Similarity graph for preinstruction (left column), postinstruction (central column), and post-4y-instruction (right column) data. The squares at the vertexes of an equilateral triangle represent the practical or everyday, descriptive and explicative typologies. The concentric circumferences mark the distances 0.25, 0.5, 0.75, and 1 from the vertexes. Blue circles represent the students belonging to the research sample, while red diamonds those belonging to the control group.

average, better than the control group, as supported by the statistically significant differences shown in Table III. Similarities in G3 graduates do not show evident variations, except for some minor changes from a pair of students slightly reducing the distances from the ex-type ideal profile. Students in G4, now graduates involved in Ph.D. studies, postgraduate specialization, or research fellowships, were among those who performed better immediately after the workshop together with few others from G3 and G2, are characterized by a clearer ex-type approach, except for two of them with increasing similarities with respect to the de-type profile. Even not considering a straightforward relationship between the percentage of students with durably high ex-type similarity index and instruction success, it seems plausible that those students who, in the last four years, got better opportunities to continue their high-level instruction and performed researchlike activities, further increased their similarity towards an ex-type profile, with respect to those who left the university to work in a company (or in any other field different from scientific research), showing a regress towards a de-type approach of problem solving.

V. CONCLUSION

The efficacy of OI-based teaching or learning strategies has been deeply investigated in the past decades [23,33–36]. Four years ago we explored this topic by quantitatively investigating the problem solving performances of a sample of mechanical engineering undergraduates who experienced an OI-based learning environment aimed at applying conceptual knowledge already acquired and promoting scientific reasoning. The global improvement in the students' problem-solving strategies, quantitatively established by the percentage increase of pre to post-instruction ex-type answers to different questions, was a clear indicator of the efficacy of OI-based learning activities to activate more efficient problem solving skills. Nevertheless, our results highlighted that OI experiences contributed positively to increase student abilities in problem solving by providing the activation of appropriate epistemological resources, quantitatively estimated by measuring the similarity index to an explicative typology of the problem solver profile [23].

In this work we have addressed the question of the lasting of student outcomes four years after having experienced an OI-based learning environment on thermal science. We have administrated the same questionnaire used in the previous study with everyday problems about thermal phenomena and applied the same research method. In addition, here a control group has been suitably selected for comparing the outcomes provided by the research sample with those adopted by students who experienced the same instruction path except for not having attended

the OI-based workshop four years ago. The differences on the average levels of similarity with the three ideal problem-solving profiles achieved by the control group with respect to the research sample are statistically significant, providing some level of trustworthiness to the results of this study.

Even considering the limits of a pilot study, a first notable result is that none of the involved students has shown a regression towards the practical or everyday epistemological profile, confirming the overall robustness of an OI-based instruction performed after a more traditional lecture-based class. This finding is particularly stable for a subsample of 7 students out of 30, namely, about 23%, where the highest values of similarity index have been recorded with respect to the explicative typology of problem solving. The stability of learning outcomes can be viewed in line with the threshold concepts theory (TCT) [54], where it is proposed that once a student has passed certain disciplinary “conceptual gateways” or “portals of understanding” there is no “way back,” i.e., the learning that has taken place is transformative and irreversible (unlikely to be forgotten) and integrative. Nevertheless, our investigations have also shown that a further 30% of students have scored at significantly higher levels in the similarity to the same ex-type epistemological profile, compared with four years' earlier data. The remaining students tried to recover some formulas from the long term memory and mechanically apply it to solve the problem, but without achieving a meaningful solution. As a matter of fact, these students faced some more difficulties to solve the proposed problems with respect to their last performances, now increasing the similarity towards the descriptive epistemological profile, but still performing better than the control group. Finally, in this paper we reported some clues that correlate our findings to the student experiences after the conclusion of their graduation instruction. We have found that Master's degree graduates continuing to practice scientific reasoning through engineering specialized studies have shown a further improvement of problem-solving skills. On the contrary, our findings also suggest that a lack of researchlike experiences could be a cause of regression towards a more descriptive epistemological profile in those students who have not completely crossed the conceptual gateways mentioned by the TCT [54].

In summary, the success of an instruction path has never been measured over such a long temporal window. This study establishes the robustness of an OI-based method of instruction in terms of a lack of regression towards a practical or everyday approach of problem solving in thermal science, confirming the efficacy of this teaching approach to support the development of long-lasting levels of problem-solving skills. The significance of our results has been supported by statistically different findings on

control group outcomes. A continuous practice of the process of scientific inquiry seems necessary to maintain or improve the reasoning abilities needed to achieve a more meaningful conceptual understanding of the physics underlying natural phenomena. In order to be confirmed, this latter claim would need a more extended study that could be the subject of a future work.

ACKNOWLEDGMENTS

This work was partially supported by MIUR and FP7 Establish Project (EU G. A. 244749). The authors would like to thank the three anonymous referees who supported the improvement of this paper with valuable comments and substantive suggestions.

-
- [1] National Research Council (NRC), *National Science Education Standards*, National Committee for Science Education Standards and Assessment (National Academy Press, Washington, DC, 1996).
- [2] National Academy of Engineering (NAE), *The Engineer of 2020: Visions of Engineering in the New Century* (National Academies Press, Washington, DC, 2004).
- [3] National Academy of Engineering (NAE), *Frontiers of Engineering 2011* (National Academies Press, Washington, DC, 2012).
- [4] National Research Council (NRC), *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Academies Press, Washington, DC, 2012).
- [5] National Research Council (NRC), *Next Generation Science Standards: For States, By States* (National Academies Press, Washington, DC, 2013).
- [6] National Research Council (NRC), *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering* (National Academies Press, Washington, DC, 2011).
- [7] National Research Council (NRC), *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning* (National Academies Press, Washington, DC, 2000).
- [8] D. Llewellyn, *Inquiry Within: Implementing Inquiry-based Science Standards* (Corwin Press, Inc., Thousand Oaks, CA, 2002).
- [9] M. Rocard, P. Csermely, D. Jorde, D. Lenzen, H. Walberg-Henriksson, and V. Hemmo, EU Research Report No. EUR 22845 (2007), https://ec.europa.eu/research/science-society/document_library/pdf_06/report-rocard-on-science-education_en.pdf.
- [10] R. A. Streveler, T. A. Litzinger, R. L. Miller, and P. S. Steif, Learning conceptual knowledge in the engineering science: overview and future research directions, *J. Eng. Educ.* **97**, 279 (2008).
- [11] E. F. Redish and K. A. Smith, Looking beyond content: Skill development for engineers, *J. Eng. Educ.* **97**, 295 (2008).
- [12] D. Hammer and A. Elby, Tapping epistemological resources for learning physics, *J. Learn. Sci.* **12**, 53 (2003).
- [13] T. J. Bing and E. F. Redish, Analyzing problem solving using math in physics: Epistemological framing via warrants, *Phys. Rev. ST Phys. Educ. Res.* **5**, 020108 (2009).
- [14] A. Gupta and A. Elby, Beyond Epistemological Deficits: Dynamic explanations of engineering students' difficulties with mathematical sense-making, *Int. J. Sci. Educ.* **33**, 2463 (2011).
- [15] E. Kuo, M. M. Hull, A. Gupta, and A. Elby, How students blend conceptual and formal mathematical reasoning in solving physics problems, *Sci. Educ.* **97**, 32 (2013).
- [16] P. Laws, A new order for mechanics, in *Proceedings Conference on Introductory Physics Course*, edited by J. Wilson (Wiley, New York, 1997), pp. 125–136.
- [17] P. W. Laws, M. C. Willis, and D. R. Sokoloff, Workshop physics and related curricula: A 25-year history of collaborative learning enhanced by computer tools for observation and analysis, *Phys. Teach.* **53**, 401 (2015).
- [18] R. K. Thornton and D. R. Sokoloff, RealTime Physics: Active learning laboratory, in *The Changing Role of Physics Departments in Modern Universities: Proceedings of International Conference on Undergraduate Physics Education*, edited by E. F. Redish and J. S. Rigden (American Institute of Physics, Woodbury, 1997), pp. 1101–1118.
- [19] R. K. Thornton and D. R. Sokoloff, Assessing student learning of Newton's laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula, *Am. J. Phys.* **66**, 338 (1998).
- [20] D. R. Sokoloff, P. W. Laws, and R. K. Thornton, Real-Time Physics: Active learning labs transforming the introductory laboratory, *Eur. J. Phys.* **28**, S83 (2007).
- [21] J. Bernhard, Insightful learning in the laboratory: Some experiences from ten years of designing and using conceptual labs, *Eur. J. Eng. Educ.* **35**, 271 (2010).
- [22] N. Pizzolato, C. Fazio, and O. R. Battaglia, Open inquiry-based learning experiences: A case study in the context of energy exchange by thermal radiation, *Eur. J. Phys.* **35**, 015024 (2014).
- [23] N. Pizzolato, C. Fazio, R. M. Sperandeo Mineo, and D. Persano Adorno, Open-inquiry driven overcoming of epistemological difficulties in engineering undergraduates: A case study in the context of thermal science, *Phys. Rev. ST Phys. Educ. Res.* **10**, 010107 (2014).
- [24] F. Marton, Describing and improving learning, in *Learning Strategies and Learning Styles*, edited by R. R. Schmeck (Plenum Press, New York, 1988), p. 53.
- [25] F. Marton and S. Booth, *Learning and Awareness* (Lawrence Erlbaum, Mahwah, NJ, 1997).

- [26] J. T. E. Richardson, The concepts and methods of phenomenographic research, *Rev. Educ. Res.* **69**, 53 (1999).
- [27] C. Fazio, O. R. Battaglia, and B. Di Paola, Investigating the quality of mental models deployed by undergraduate engineering students in creating explanations: The case of thermally activated phenomena, *Phys. Rev. ST Phys. Educ. Res.* **9**, 020101 (2013).
- [28] G. E. Francis, J. P. Adams, and E. J. Noonan, Do they stay fixed?, *Phys. Teach.* **36**, 488 (1998).
- [29] J. Bernhard, Does active engagement curricula give long-lived conceptual understanding?, *Physics Teacher Education Beyond 2000*, edited by R. Pinto and S. Surinach (Elsevier, Paris, 2001), pp. 9–752.
- [30] M. D. Herron, The nature of scientific enquiry, *School Rev.* **79**, 171 (1971).
- [31] R. W. Bybee, An instructional model for science education, *Developing Biological Literacy* (Biological Sciences Curriculum Study, Colorado Springs, CO, 1993).
- [32] H. Banchi and R. Bell, The Many Levels of Inquiry, *Sci. Child.* **46**, 26 (2008).
- [33] C. A. R. Berg, V. C. B. Bergendahl, and B. K. S. Lundberg, Benefiting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus an open-inquiry version of the same experiment, *Int. J. Sci. Educ.* **25**, 351 (2003).
- [34] R. A. Krystyniak and H. W. Heikkinen, Analysis of verbal interactions during an extended, open-inquiry general chemistry laboratory investigation, *J. Res. Sci. Teach.* **44**, 1160 (2007).
- [35] M. Zion, M. Slezak, D. Shapira, E. Link, N. Bashan, M. Brumer, T. Orian, R. Nussinowitz, D. Court, B. Agrest, R. Mendelovici, and N. Valanides, Dynamic, open inquiry in biology learning, *Sci. Educ.* **88**, 728 (2004).
- [36] I. Sadeh and M. Zion, The development of dynamic inquiry performances within an open inquiry setting: A comparison to guided inquiry setting, *J. Res. Sci. Teach.* **46**, 1137 (2009).
- [37] N. Trautmann, J. MaKinster, and L. Avery, What makes inquiry so hard? (and why is it worth it), *Proceeding at the Annual Meeting of the National Association for Research in Science Teaching*, (Vancouver, BC, Canada, 2004).
- [38] C. Quintana, X. Zhang, and J. Krajcik, A framework for supporting meta cognitive aspects of on-line inquiry through software-based scaffolding, *Educ. Psychol.* **40**, 235 (2005).
- [39] D. Persano Adorno and N. Pizzolato, An inquiry-based approach to the Franck-Hertz experiment, *Il Nuovo Cimento C* **38**, 109 (2015).
- [40] R. S. Schwartz, N. G. Lederman, and B. A. Crawford, Developing views of nature of science in an authentic context: an explicit approach to bridging the gap between nature of science and scientific inquiry, *Sci. Educ.* **88**, 610 (2004).
- [41] L. Flick and N. G. Lederman, Scientific Inquiry and Nature of Science, in *Contemporary Trends and Issues in Science Education*, edited by L. Flick and N. G. Lederman (Springer, New York, 2006), Vol. 25, ISBN 978-1-4020-2671-3.
- [42] C. K. Capps and B. A. Crawford, Inquiry-based instruction and teaching about nature of science: are they happening?, *J. Sci. Teach. Educ.* **24**, 497 (2013).
- [43] G. M. Quan and A. Elby, Connecting self-efficacy and views about the nature of science in undergraduate research experiences, *Phys. Rev. Phys. Educ. Res.* **12**, 020140 (2016).
- [44] B. K. Hofer and P. R. Pintrich, *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (Lawrence Erlbaum, Mahwah, NJ, 2002).
- [45] D. Hammer and A. Elby, Tapping epistemological resources for learning physics, *J. Learn. Sci.* **12**, 53 (2003).
- [46] P. G. Jasien and G. E. Oberem, Understanding of elementary concepts in heat and temperature among college students and K-12 teachers, *J. Chem. Educ.* **79**, 889 (2002).
- [47] R. A. Streveler, B. M. Olds, R. L. Miller, and M. A. Nelson, Using a Delphi study to identify the most difficult concepts for students to master in thermal and transport science, in *Proceedings of ASEE Annual Conference* (American Society for Engineering Education, Washington DC, 2003).
- [48] E. F. Redish and K. A. Smith, Looking beyond content: Skill development for engineers, *J. Eng. Educ.* **97**, 295 (2008).
- [49] R. L. Miller, R. A. Streveler, D. Yang, and A. I. Santiago-Román, Identifying and repairing student misconceptions in thermal and transport science: Concept Inventories and Schema Training Studies, *Chem. Eng. Educ.* **45**, 203 (2011).
- [50] M. Prince, M. Vigeant, and K. Nottis, Development of the Heat and Energy Concept Inventory: Preliminary results on the prevalence and persistence of engineering students' misconceptions, *J. Eng. Educ.* **101**, 412 (2012).
- [51] A. Fernández and S. Gómez, Solving non-uniqueness in agglomerative hierarchical clustering using multidendrograms, *J. Classif.* **25**, 43 (2008).
- [52] R. Gras, F. Suzuki, F. Guillet, and F. Spagnolo, *Statistical Implicative Analysis: Theory and Applications* (Springer, New York, 2008).
- [53] A. Markos, G. Menexes, and I. Papadimitriou, in *Classification as a Tool for Research: Proceedings of the 11th International Federation of Classification Conference*, edited by H. Loracek-Junge and C. Weihs (Springer, Berlin, Germany, 2010), p. 409.
- [54] J. H. F. Meyer and R. Land, Threshold concepts and troublesome knowledge (2): Epistemological considerations and a conceptual framework for teaching and learning, *Higher Educ.* **49**, 373 (2005).