MEasurement in STEM Education
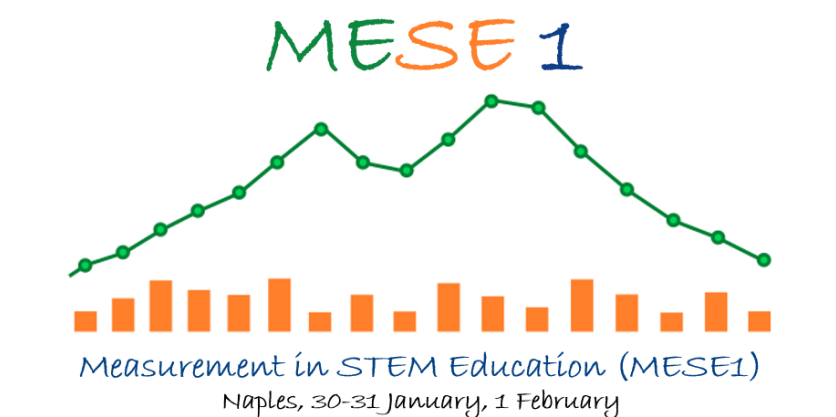
MESE 1

Measurement in STEM Education (MESE1)

Naples, 30-31 January, 1 February

DISP Dipartimento di Scienze Politiche Università degli studi di Napoli Federico II

Università degli Studi di Napoli Federico II · Dipartimento di FISICA "ETTORE PANCINI"

Report of Contributions

Contribution ID : **2**   Type : **Oral Contribution**

Investigation of Systems Thinking Skills of students aged 11 to 14 years old

lunedì 30 gennaio 2023 16:35 (20)

The education of responsible citizens, capable of becoming active members of society, is one of the goals of the European Commission (Hazelkorn, 2015). It aims to create conditions for science education that foster students' understanding of the complexity of the world of which they are a part. Green Comp, a framework developed by the European Commission (Bianchi et al., 2022) recognizes Systems Thinking as one of twelve competencies to increase and develop the knowledge and attitudes to work, live and act in a sustainable world. Systems Thinking is, therefore, an approach that enables students, regardless of their future careers, to develop the knowledge and skills to consciously take part in local and global challenges. As part of my research project, I, therefore, administered the Systems Thinking Assessment Italia (STAI) test to students in Lower Secondary School in the Province of Trento to investigate the Systems Thinking competencies of students aged 11 to 14. The STAI test is a translation and adaptation from Greek Cypriot to Italian of the Systems Thinking Assessment (STA) test (Κωνσταντινίδη, 2015). It consists of 27 questions divided into four categories (elements in the system, interactions in the system, flows in the system, dynamics in the system). Each category is further divided into skills, the Systemic Thinker skills. The analysis covers a sample of 709 students from nine Secondary Schools (240 students in grade one, 232 students in grade two, 237 students in grade three). Rasch analysis, ANOVA on the sample, ANOVA on the test, and a frequency analysis show that the test in its entirety is appropriate for this age group but is unable to discriminate the identified categories. Students show basic Systems Thinking skills that do not improve significantly over the three-year period, confirming the lack of a program to develop these skills. This investigation is part of a doctoral project in its concluding stages entitled 'Systems Thinking in the Lower Secondary School: field survey in the Province of Trento and intervention study'.

Research Strand

Methods and Methodological Aspects in Science Education Research

Primary author(s) :   ZANELLA, Sara (Free University of Bozen)

Presenter(s) :   ZANELLA, Sara (Free University of Bozen)

Session Classification :   Measurement of cognitive and affective variables in STEM (Parallel Contributed Symposium)

Track Classification :   Session 2a – Measurement of cognitive and affective variables in STEM

Contribution ID : **3**   Type : **Oral Contribution**

Combining qualitative and quantitative analyses to investigate the role of virtual simulations in the evolution of primary school students' mental models about electrostatics

lunedì 30 gennaio 2023 16:35 (20)

We combined qualitative and quantitative analyses to investigate the role of virtual simulations in the evolution of primary school students' mental models about electrostatics with respect to the target school scientific model of microscopic charges by analyzing the answers of two groups of 9/10 years old pupils inside instructional sequences that combined two hands-on real (R) activities using balloons and jackets with the same two virtual activities (V) using PhET simulations. Five classes in three schools (a total of n = 83 students, 43 males and 40 females) were chosen to carry out the "first-real" (RV) sequence while five classes in three other schools (a total of n = 85 students, 41 males and 44 females) carried out the "first-virtual" (VR) sequence. The two groups were organized so as to be homogeneous from an ethnic, cultural, social and economic viewpoint, ensuring that the typology of students would not affect the results. Data collection was made through pre-observation and post-observation worksheets for each real and virtual activity concerning a descriptive dimension as well as an explanatory dimension, leading to a total eight worksheets for each student. As for the descriptive dimension, related to the multiple-choice question "What happens? Choose a description for the physical behavior of the two objects", students' answers were classified with respect to three categories: "Concordant", "Discordant" and "No Answer". The explanatory dimension, investigated through the open question "Why does it happen? Justify your choice", required a qualitative and interpretative analysis of the texts and of the graphic representations proposed by pupils. To classify students' answers and establish a progression with respect to their degree of adequacy to the target model, we defined five categories associated to five levels from 0 (lowest adequacy) to 4 (highest adequacy) following an approach based on the empirical learning progressions. Quantitative data analysis was applied to compare the efficacy of the RV and VR sequences. In the case of the descriptive dimension, we applied the chi-square test since our aim was to estimate whether the difference between the percentages of concordant, discordant and not given answers between the two sequences was statistically significant. In the case of the explanatory dimension, where we wanted to test the statistical significance of the difference in the ranking of students' levels from 0 to 5 between the two sequences, we applied the Mann-Whitney U-test. In both cases, the null-hypothesis was that the differences between the RV and the VR sequence were not statistically significant: the difference between the two sequences was considered as statistically meaningful when the probability that the null hypothesis was true (p-value) resulted lower than 5%. This mixed qualitative and quantitative approach allowed to show that virtual simulations do improve the level of adequacy of students' answers to the target model, but this improvement is not transferred to new phenomena, so that the capability of relating the developed model to the real world remained the same at the end of both sequences.

Research Strand

Primary author(s) : BOZZO, Giacomo (Department of Biology, Ecology and Earth Science. University of Calabria (UNICAL), Cosenza, Italy)

Co-author(s) : LOPEZ, Victor (Centre for Research of Science and Mathematics Education (CRECIM), Autonomous University of Barcelona (UAB) – Spain); COUSO, Digna (Centre for Research of Science and Mathematics Education (CRECIM), Autonomous University of Barcelona (UAB) – Spain); MONTI, Francesca (Department of Computer Science, University of Verona (UNIVR), Verona, Italy)

Presenter(s) : BOZZO, Giacomo (Department of Biology, Ecology and Earth Science. University of Calabria (UNICAL), Cosenza, Italy)

Session Classification : Assessment of Instructional Strategies in STEM Education (Parallel Contributed Symposium)

Track Classification : Session 2b – Assessment of Instructional Strategies in STEM Education

Contribution ID : 4   Type : **Oral Contribution**

Investigating through worksheets the impact of short formative modules on prospective teachers' didactic projects about specific topics: qualitative analysis and data visualization.

lunedì 30 gennaio 2023 17:15 (20)

We describe how we approached a study on a group of 65 primary school prospective teachers (PPTs) aimed at investigating through post-assessment questions, worksheets and final interviews how they would project a learning path on the subject of "heat, temperature and energy" after following a short didactic module involving the use of infrared thermal cameras and temperature online sensors(1). The intervention was proposed inside the laboratory part of the Physics Education course of the combined bachelor and master degree in Primary School education of the Italian University of Verona. The module was organized into four phases. The first and third phases were developed as interactive experiment-based lectures following active learning strategies to introduce the operational definition of temperature, the concept of energy, the use of the online sensors and the use of the thermal cameras. In the second and fourth lab-work phases, four experiments were autonomously carried out by the PPTs divided into groups. Data collection was made at the end of the course. In the post-assessment questions, PPTs were asked to list the ways they know for measuring temperature. In the worksheets, they were asked to plan a didactic project for their pupils about thermal phenomena, highlighting the addressed concepts and the corresponding related activities. Final interviews were mainly focused on whether and how PPTs used or would use a thermal camera in their proposed activities and, if so, related to which concepts. The investigation on how much the formative module and methodologies did help PPTs in understanding the proposed concepts was made through the following research questions: 1) Do PPTs cite (in the proposed post-assessment question) thermal cameras among the possible instruments for measuring temperature? 2) Which concepts and corresponding activities do PPTs choose (in the worksheets) in constructing their didactic project on thermal phenomena? 3) Do (from the worksheets) or would (from the final interviews) PPTs propose the use of thermal cameras in their activities and, if so, related to what concepts? Data analysis was performed following an iterative process of Qualitative Analysis(2) by identifying directly from students' answers a set of categories and refining them through successive re-readings of students' reports. Attention was paid: a) on the instruments (online sensors and thermal cameras) as cited in the post-assessment questions and as utilized in the activities proposed in their planned learning paths; b) on the concepts addressed by PPTs as related to the instruments utilized in the activities. The frequencies and percentages of the categorized answers were visualized as bar charts. We will present some examples of the answers given by students and a description of the process that lead us to the final categorization and interpretation of the obtained results.

(1)Monti and Daffara (2021) Journal of Physics Conference Series 1929(1)012020 (2)Miles,Huberman and Saldaña (2014) "Qualitative data analysis: a methods sourcebook (Third Edition)". Thousand Oaks, California: Sage Publications, Inc.

Research Strand

Assessment and Evaluation in science learning processes

Primary author(s) :   MONTI, Francesca (University of Verona); Prof. DAFFARA, Claudia (Univer-

sity of Verona)

Presenter(s) :   MONTI, Francesca (University of Verona)

Session Classification :   Assessment of Instructional Strategies in STEM Education (Parallel Contributed Symposium)

Track Classification :   Session 2b – Assessment of Instructional Strategies in STEM Education

Contribution ID : **6**   Type : **Oral Contribution**

Peer learning in higher education: An effective response to the university students' dropout problem

lunedì 30 gennaio 2023 16:55 (20)

The process of globalisation has shown that countries can play a key role on the international stage provided that their citizens are given the opportunity to build a significant background in science, technology, engineering and mathematics (STEM). One of the critical issues regarding the low number of degrees in STEM faculties is the considerable dropout rate in the first years of higher education. In this context, Italy is no exception with about 20% of its university students who drop out within the first two academic years. In a recent study based on administrative data from the Italian Politecnico di Milano, it has been pointed out that the most important factor which allows to predict Politecnico di Milano student dropout is the number of university educational credits gained across the first term of the first academic year. Since active methods employed in academic courses appear to enhance students' learning more than traditional lectures even in the context of large size classes, their use could lead to a decrease in the dropout rate. In the academic year 2021-2022 we carried out a case study which involved about two hundred freshmen attending the "Fisica Sperimentale A+B" course at Politecnico di Milano. In addition to traditional lectures and drills, seven peer learning sessions were offered to these students. During each peer learning session, the learners answered a questionnaire consisting of three multiple choice items based on some Physics topics and whose provision was implemented by using the students' response system Socrative. Immediately after the questionnaire, freshmen in the classroom would discuss the quizzes in small groups for few minutes. At the end of this debate, they retook the same questionnaire. Finally, the instructor briefly illustrated the correct as well as the incorrect alternatives of each item and the percentage of answers ascribed to each possible option were shown to the students. Considering that some students attended that course in person while others attended the lessons on line, we had both an experimental and a control group. Their initial knowledge in Physics was checked and compared through a questionnaire. This instrument was administered at the beginning of the academic course and based on multiple choice items on some Physics topics which were completely different from the ones employed during the peer learning sessions. In order to evaluate the effectiveness of this education methodology, we examined the freshmen's achievement in their Physics course final examination during the first exam session. Our findings show that the success rate of the experimental group was higher than the control group and this difference was statistically significant. Moreover, the calculated effect size highlighted that the association between the final examination pass rate and the peer learning sessions attendance was relatively strong. Furthermore, we investigated the possible correlation between the final examination pass rate and the number of peer learning sessions attended by the students. On balance, these results appear to confirm that our innovative educational methodology may be effective in the mitigation of the university students' dropout rate.

Research Strand

University Education

Primary author(s) :   BOZZI, Matteo (Politecnico di Milano, Department of Physics, Milan, Italy)

Co-author(s) :   Mr. MAZZOLA, Roberto (Politecnico di Milano, Department of Physics, Milan, Italy); Prof. ZANI, Maurizio (Politecnico di Milano, Department of Physics, Milan, Italy)

Presenter(s) :   BOZZI, Matteo (Politecnico di Milano, Department of Physics, Milan, Italy)

Session Classification :   Assessment of Instructional Strategies in STEM Education (Parallel Contributed Symposium)

Track Classification :   Session 2b – Assessment of Instructional Strategies in STEM Education

Contribution ID : **7**   Type : **Oral Contribution**

An extensive questionnaire about emergency remote teaching: more than 3000 engineering students respond about their perceptions on online didactic activities.

lunedì 30 gennaio 2023 16:15 (20)

By 11 March 2020, the phrase "COVID-19" had officially entered everyday life across most of the word. Each level of education suddenly faced new changes and new challenges. Emergency remote teaching became widespread, and new methodologies to deliver classes and courses were adopted by educational institutions. In February 2020, the Politecnico di Milano introduced a series of focused and systemic actions in order to support the passage to completely online teaching and to ensure the continuity of the activities that were previously developed in the classroom. In our work, we focus on the impact of the remote learning experience on STEM disciplines, by analysing the perceptions of engineering students enrolled at the Politecnico di Milano. The subjects were recruited from all engineering courses, from the first to the fifth year, and were asked to complete a multidimensional survey. The questionnaire was proposed in July 2021 at the end of the second semester, referring to the didactic activities held in the 2019-2020 and 2020-2021 academic years. More than 3000 students completed the entire survey, by answering the questions that were proposed in a Likert scale, from 1 to 5. This large sample was composed of 66% male and 33% female students. Almost the 70% of the students were attending Bachelor's degree courses, while the remaining were attending Master's degree courses. The survey was composed of 66 questions regarding the perceptions and the challenges of online education, compared with the "state of the art" before COVID-19, divided into 6 main groups: Remote Teaching, Subjective Well Being, Metacognition, Self-Efficacy, Identity and Socio-Demographic information. In this work we described the entire survey and we focused on the items concerning Metacognition and Self Efficacy but we also gave a first glance to the results concerning the remote teaching section. We performed preliminary analysis, by computing frequency distribution and descriptive statistics concerning the remote teaching section, then, by using Cronbach's alpha test, confirmatory factor analysis and the t-test, we performed a more in-depth analysis concerning the outcomes of metacognition and self-efficacy. Data analysis shows that students clearly appreciated how the Politecnico di Milano dealt with the organization of the mandatory online courses, but in the same time the results indicates that they complained about their relationships with classmates during remote teaching. Data also suggests an improving in effective learning strategies performed by students. We can say that students, in some sense, overcame the difficulties due to the emergency remote teaching by improving their cognitive processes. Potentially, this last result could have an important impact in teaching methodologies, adding a point in favour of distance learning. But the topic is so sensitive and discussed that it surely needs to be investigated in more details.

Research Strand

University Education

Primary author(s) : MAZZOLA, Roberto (Department of Physics, Politecnico di Milano)

Co-author(s) :   Dr. BOZZI, Matteo (Department of Physics, Politecnico di Milano); Prof. ZANI, Maurizio (Department of Physics, Politecnico di Milano)

Presenter(s) :   MAZZOLA, Roberto (Department of Physics, Politecnico di Milano)

Session Classification :   Measurement of cognitive and affective variables in STEM (Parallel Contributed Symposium)

Track Classification :   Session 2a – Measurement of cognitive and affective variables in STEM

Contribution ID : **8**   Type : **Oral Contribution**

Measuring the impact of changing the formulation of a mathematical task

lunedì 30 gennaio 2023 12:00 (15)

Learning achievements are usually assessed by administering tasks to students- ranging from individual assessment to Large-scale surveys. Task design and administration are key issues in an assessment process. In the last decades, quantitative research investigated systemic changes in assessment modality, such as how the switching from a paper-and pencil administration of a task to a computer-based testing impacts on students' answers, or from static to interactive administration. Less research has been done at item level on the impact of changes in the formulation of a task. Students' answers to an item depend on a very complex system of relationships between the elements of the formulation, personal features, teachers' style, and the learning noosphere. It is a somehow chaotic system where no direct cause-and-effect relationships are usually clearly identifiable. A classical research trend in mathematics education is how the linguistic formulation of a word problem impact students' answers. This issue, of course, has many significant consequences, for instance, on how items administered in international surveys are translated into the different languages. Qualitative research investigated how a change in the formulation of a problem may actually change the problem itself, for a particular student. Several theoretical frameworks are available, based on qualitative evidence, but measurement is rarely used in Mathematics Education studies, since the optimal situation is difficult to achieve. It is of course impossible to administer two different versions of the same task to the same student, since he/she should "forget" to have answered to that task the first time, before answering to the second version. To overcome such a limit, we developed and administered four anchored Mathematics achievement tests, consisting of two sections. The first one was given by a core set (named Core Test) of common items (used to anchor tests) and representing the entire achievement test from both a statistical and mathematical point of view (i.e., both in terms of estimated parameters and mathematics items' content). Compared with previous studies in Mathematics Education, our "Core Test" consists of items from Large-scale assessment, each of which administered by INVALSI to the entire Italian students' population in previous main studies. Results from data analysis confirmed the "robustness" of our Core Test that can thus serve both as a benchmark and as a definition of the latent trait. This measurement step can be inserted both in explanatory and in exploratory plans. The second section of each achievement test was given by alternative formulations of the same stem-items: each stem-item was modified by performing on it a single, well-individuated variation, thus resulting in a quasi-experimental design which can detect emerging phenomena due to changes in formulation (hence overlapping this difficulty at least at a systemic level). We present examples of application of this methodological strategy, showing how it may contribute to better define the summative aspects of assessment, help in interpreting students' behaviors when answering to an assessment question, thus giving a contribution to formative assessment, and help task-assessment designers.

Research Strand

Methods and Methodological Aspects in Science Education Research

Primary author(s) : BOLONDI, Giorgio (Free University of Bozen-Bolzano)

Co-author(s) : CASCELLA, Clelia (INVALSI); Dr. GIBERTI, Chiara (University of Bergamo); Dr. BRANCHETTI, Laura (University of Milan)

Presenter(s) : BOLONDI, Giorgio (Free University of Bozen-Bolzano)

Session Classification : Issues in Assessing STEM and Math Education (Contributed Symposium)

Contribution ID : **9**   Type : **Oral Contribution**

Why does Integrated STEM education not fit: Pinpointing the breaking points.

lunedì 30 gennaio 2023 11:00 (20)

Despite of countless financial resources at the European and Italian national level, the availability of training and many best practices as teaching resources, STEM education struggles to be integrated in the school system in holistic way. Starting from a reflection on the key points and actors involved with STEM practices in schools this contribution intends to highlight the critical points in the educational system that may hinder the process of their integration. A minimalist version of STEM integration would require the combination of at least two STEM subject areas STEM integration can also extend beyond STEM disciplines to incorporate art or any other subject and the interest and development on interdisciplinary STEM education is increasing. STEM integration could be represented as a gradient, which at one extreme is defined by a multidisciplinary approach of two or more disciplines taught and learned separately but within a unifying theme, and at the other end a transdisciplinary approach, the learning goals of the different disciplines are designed through a learning scenario based on real-world problems. Yoder and colleagues define six key elements of a student-centred and evidence-based educational STEM ecosystem including: effective teachers, high standards and quality curricula, formal and informal learning of STEM subjects, integrated use of technology. Aside for teaching resources, teachers need to be trained in STEM pedagogical design, implementation, and evaluation. To improve teachers' self-efficacy in teaching STEM professional development programs also need to address teachers' needs for instructional practices and provide opportunities for peer collaboration. The Scientix report on educational practices in Europe disposes a different picture. The responses show that for STEM subjects, the frontal and transmissive lecture remains the most reported pedagogical approach and most STEM teachers interviewed have not undergone any professional training of STEM subject-related methodologies or technologies. In daily practice also the lack of time required for instructional design and pedagogical implementation and a formative-oriented integrated STEM approach which does not always fit easily into the curriculum. To overcome these barriers a school curriculum ready to promote an integrated STEM approach and a shift to distributed leadership to support and promote STEM practices and teaching at the whole-school level could enable teacher collaboration and stakeholder engagement in STEM education. The school leadership should support and promote a professional development with collective and active participation, consistent with the guidelines and curriculum of an extended duration to ensure accompaniment in the application of new skills in teaching practice; as the active collaboration with the territory, like industry or museums, in classroom activities, to contextualise the learning content to a real-world application. Policy makers should encourage professional development of teachers and school leaders, as foster the introduction of also student-centred methodologies and formative assessment in learning.

Research Strand

Network Analysis in Education and Learning Processes

Primary author(s) :   NIEWINT-GORI, Jessica (Indire)

Presenter(s) : NIEWINT-GORI, Jessica (Indire)

Session Classification : Issues in Assessing STEM and Math Education (Contributed Symposium)

Contribution ID : **10**   Type : **Oral Contribution**

Artificial Neural Network Analysis of Physics and Engineering freshmen FMCE responses

lunedì 30 gennaio 2023 16:15 (20)

We administered FMCE concept inventory (Thornton and Sokolof, 1998) at the beginning of college studies to young Physics and Engineering freshmen of the University of Trieste during the Academic Year 2021- 2022. We wanted to spotlight if there were influences on the Physics knowledge acquired involved by the context of instruction (Bao and Redish, 2006). We provided the Italian translation of the Concept Inventory to administer it to freshmen. The translation from the original English FMCE was performed using a conceptual translation model. To investigate conceptual coherence, we analysed the responses exploring if there was knowledge fragmentation in using the concepts, thus looking for lack of conceptual change (diSessa, 1993, 2014; Vosniadou, 1994). We conducted two different kinds of analysis. The first one was based on multivariate descriptive statistics. Then we undertook an in-depth analysis based on the Artificial Neural Network (ANN) method (Lamb et al., 2014). This versatile framework (Amoo et al., 2018) computes relationships based on the interaction of multiple connected processing elements (Lamb et al., 2014; Pinkus, 1999). We used ANN as an analytical data method to investigate possible statistical evidence of what descriptive analysis suggested. In ANN design, we adopted the Sanger rule, also known as Sequential Principal Components Analysis, which forces neurons to represent a well-ordered set of principal components of the data set (Sanger, 1989). Artificial Neural Network analysis supplied a deeper insight into data-set statistical correlation. The Artificial Neural Network analysis validated what was depicted by descriptive statistical analysis. This gave us a noteworthy method to feature students' conceptual knowledge expressly about force and motion topics, even if we need to observe a more representative sample for a more robust statistical inference. One relevant result concerns the responses clustering according to conceptual misunderstanding (Hammer, 1998). This shows that most of the sampled freshmen begin their College studies in Physics and Engineering still knowing force and motion concepts intuitive or naïve. Only a small group of students, mainly belonging to the Physics freshmen group, exhibit a robust conceptual building in the description of force/motion phenomena. This clearly occurs by probing the neural network and relying on how different groupings of students follow the answers identified. Exploring the correlation between answers and past-curriculum, we observed that a long-exposure to physics curricular studies enacts a better conceptual understanding, as would be. What remarkable occurs is the negative trend in students with a three-year past-curriculum. Even if they experienced Physics studies at a heightened performance of cognitive development (mostly between 17-19 years old), they would suffer the lack of conceptual change in force/motion description. The difficulties are steady in students with only early two-year exposition to Physics studies. Finally, by present-curriculum choice, we confirmed the differences between students' group affiliations. Students from the Civil, Environment, Electronic and Informatics Engineering group present more conceptual difficulties than Physicists and Industrial Naval Engineers.

Research Strand

Assessment and Evaluation in science learning processes

Primary author(s) :   BOLOGNA, Valentina (Physics Department, University of Trieste)

Co-author(s) :   DIAZ-GUERRA, Fabrizio (Physics Department, University of Trieste);   LONGO, Francesco (Physics Department, University of Trieste)

Presenter(s) :   BOLOGNA, Valentina (Physics Department, University of Trieste)

Session Classification :   Assessment of Instructional Strategies in STEM Education (Parallel Contributed Symposium)

Track Classification :   Session 2a – Measurement of cognitive and affective variables in STEM

Contribution ID : **11**   Type : **Oral Contribution**

Spatial abilities and the gender gap in mathematics (Matabì)

lunedì 30 gennaio 2023 11:40 (20)

Learning mathematics is a general problem in Italy, but it is more relevant for females. Male students typically outperform their female classmates in maths test scores from the earliest years of schooling, and the gap worsens as school grades progress.

The different experiences and expectations about potential abilities of boys and girls are key to understanding the possible causes of the observed relative differences in education performance. Early life experiences are essential for developing a child's cognitive capacities, and environmentally induced gender differences may arise when children are exposed to heterogeneous sources of development opportunities. One of the sources of gender inequalities in mathematics is related to differences in the acquisition of visuospatial abilities between girls and boys from a very young age.

Existing works find that playing with specific toys facilitates learning in maths and science. Boys usually gain more experience than girls because of different parental and educators' beliefs and behaviour regarding the gender-specific suitability of toys. Moreover, self-confidence and anxiety are not gender-neutral and can affect educational outcomes.

The Matabì project aims to enhance spatial abilities and reduce the gender gap via construction play (i.e., by using building toys and Lego Duplo brick sets). Using the bricks should help students process abstract concepts, while the playful approach should reduce maths anxiety. Girls should benefit more from Matabì than males, who are (on average) able to develop these skills early on in life.

The project started in October 2022, so we would present the overall structure and what has been accomplished so far. This should be engaging in three respects: the teachers' reinforcement of their spatial abilities, the instruments used to collect the outcomes, and the impact evaluation design conducted with Randomized Control Trial (RCT).

Teachers' training includes a pre- and post-standardised test, the mental rotation section of the Revised Purdue Spatial Visualization Test. The RCT involves around 60 classes in five schools in Torino (60 teachers and 1200 third and fourth grade pupils). Within each school, we randomly assigned a third of the teachers to the control group and two-thirds to the treatment group.

The study will provide evidence of the effects of teaching methods exploiting construction play. The primary research questions are:

1. What is the impact on the spatial abilities of teachers?

2. What is the impact on the spatial abilities of pupils?

3. How does the impact differ by gender?

4. How does the treatment affect the gender gap in mathematics?

Research Strand

Primary author(s) :   ROMANO, Barbara (Fondazione Giovanni Agnelli)

Co-author(s) :   BALLATORE, Maria Giulia (Politecnico di Torino); DI LIBERTO, Adriana (Università di Cagliari); GENOVA, Chiara (Fondazione Giovanni Agnelli); GIUA, Ludovica (Università di Cagliari); TABACCO, Anita (Politecnico di Torino)

Presenter(s) :   ROMANO, Barbara (Fondazione Giovanni Agnelli)

Session Classification :   Issues in Assessing STEM and Math Education (Contributed Symposium)

Contribution ID : **12**   Type : **Invited talk**

An application of Differential Item and Bundle Functioning analysis to the study of gender differences in Mathematics Education: implications for educational practitioners.

martedì 31 gennaio 2023 11:30 (25)

The study of gender differences in mathematics attainment has exponentially increased over time as closing the gap between boys and girls is a priority for both research and policy. Mathematics has in fact been listed by the European Commission as one of the 'key competences' necessary to all for personal fulfilment and development. Nonetheless, there are still several countries around the world in which boys significantly outperform girls in mathematics with negative implications at both individual and collective level (i.e., in terms of expected employability and wage as well as in terms of societal and economic development, especially considering that the demand of STEM-related job is going to increase). The current research builds on a Marie Curie project aimed at exploring the possible association between gender differences (particularly, female underachievement) in mathematics and environmental socio-cultural and economic factors. Building on results showing that the more traditional the 'field' in which students grow their gender identity, the better the boys' attainment in mathematics compared with girls, the current study goes a step further by showing how measurement can be used to help educational practitioners in understanding the relationship between students' gender and their attainment in mathematics. Within the framework of the Rasch analysis, educational research has frequently employed Differential Item Functioning (DIF) analysis to measure the (possible) association between the probability of successfully encountering each single (mathematics) item and a single students' characteristic (such as gender). Differential Functioning occurs when examinees from distinct groups, but matched on ability, have different probabilities of answering an item correctly. Differential functioning can relate to a single item or to a bundle of items. The former is referred as Differential Item Functioning (DIF), the latter as Differential Bundle Functioning (DBF). In the current study, after showing results from a systematic literature review about the employment of DBF in educational research, I present an application of both DIF and DBF to Large-Scale Assessment (LSA) data, collected in Italy by the Italian national institute for the evaluation of educational system, in 2017, at Grade 10 (on average, 15-years old students). Results were used to discuss (i) strengths and weaknesses of both DIF and DBF; (ii) similarities and differences of these two analyses and, thus, how each of them can be employed to support educational practitioners in dealing with gender differences; as well as (iii) the use of LSA data in Mathematics Education research.

Research Strand

Primary author(s) :   CASCELLA, Clelia (INVALSI)

Presenter(s) :   CASCELLA, Clelia (INVALSI)

Session Classification :   Assessing student competence: current methodological issues in

case studies (Invited Symposium)

Track Classification : Assessing student competence: current methodological issues in case
studies (Invited Symposium)

Contribution ID : **13**   Type : **Oral Contribution**

Mathematics achievement at the end of upper secondary school during COVID-19 pandemic: insights from the INVALSI national assessment

martedì 31 gennaio 2023 11:05 (25)

Education stakeholders in times of the Covid-19 pandemic require a broad empirical base to recover from the crisis and to strengthen the resilience of education systems in the future. National assessments of students learning outcomes are potentially useful sources of education data during the COVID-19 pandemic, playing a pivotal role in monitoring school systems and improving education quality. In Italy, the National Institute for the evaluation of the education and training educational system (INVALSI) every year carries out a standardized national assessment of students' achievements in primary and secondary education. From the school year (SY) 2018-2019, the INVALSI national testing program has been extended to the last year of upper secondary school, thus allowing to depict of an overall picture of students' mathematics achievement at a key point of transition to tertiary education and employment. Results are reported not only in terms of numerical scores but also as proficiency levels, to offer substantial information on the proficiency status at the system level. The attribution of an explicitly described level is also supposed to allow the students, their families, and the teachers to have more significant and useful feedback compared to a simple score, thanks to the direct link to the content area covered by the test. The present work aims at providing an overall picture of students' performance in the INVALSI mathematical assessment at the end of upper secondary school in the school year 2020-21, about one year after the COVID-19 outbreak in Italy, taking into account their pre-pandemic mathematics proficiency level (Grade 10, s.y. 2017-18). Protective factors and possible sources of inequalities in achievements during the COVID-19 crisis are also explored, by adopting a multilevel approach. The overall differences between pre-pandemic and pandemic cohorts in Italy suggest a pandemic achievement gap in mathematics at the end of upper secondary school, with a higher number of students resulting as low performers, i.e. "who are likely to use basic skills and procedures mainly acquired in lower secondary school and, partly, at the end of the first two years of upper secondary school [...]" or maximum "knowing the basic mathematical concepts as outlined in the national guidelines for mathematics in the first two years of upper secondary school" [...]. Considering their pre-pandemic starting points (G10) in a retrospective perspective, it emerges that although most of the G13 low performers were already struggling with mathematics, about one out of three moved from intermediate-high (G10 scale) to lowest levels (G13 scale). More encouraging results are those from another subgroup of pandemic-cohort students who maintained intermediate to high performance in mathematics (with respect to G10 and G13 scales), suggesting positive patterns of adaptation in the context of adversity due to the COVID-19 crisis. Multilevel analyses provide further insights into the relevance of different variables in supporting the students' relative progresses in mathematics during the pandemic, both at the individual and the contextual level.

Research Strand

Primary author(s) : DESIMONI, Marta (INVALSI); NOLLI, Nicoletta; LASORSA, Cristina (IN-VALSI); CERAVOLO, Rosalba (INVALSI); MATTEI, Alessia (INVALSI)

Presenter(s) : DESIMONI, Marta (INVALSI)

Session Classification : Assessing student competence: current methodological issues in case studies (Invited Symposium)

Track Classification : Assessing student competence: current methodological issues in case studies (Invited Symposium)

Contribution ID : **14**   Type : **Invited talk**

Clustering educational data: a high school students' performance analysis

martedì 31 gennaio 2023 11:55 (25)

In this talk, we first briefly discuss the definitions of Educational Data Mining and Learning Analytics, providing an overview of the most used statistical methods in both fields, as well as a synopsis of their differences and similarities. The possible aims of the different methods, aimed at uncovering hidden patterns in educational data, are stressed pointing to the possible services provided to the whole school community. In the following, we focus on clustering methods for educational data: we start from traditional methods (hierarchical, partitive, and density-based), we proceed with dimension reduction methods, such as factorial k-means and reduced k-means, and we present some methods which incorporate the longitudinal dimension. Then, we present a pilot analysis carried out on a dataset reporting the performance of a class of high school students in three periods (which were treated as three separate datasets), using hierarchical clustering, partitive clustering (k-means), factorial k-means and reduced k-means techniques. The goal of the analysis is to show how the composition of groups and the number of groups vary in each period and which are the factors that influence the creation of groups. The partitions obtained with these algorithms were compared in terms of reliability using the average silhouette width index. Reduced k-means and k-means generated similar results and we can say that these results were the most acceptable considering the average silhouette width. Hierarchical clustering generated the same results as the former algorithms only in the first two periods of time. The results generated by factorial k-means differ from the other methods and as suggested by the values of the average silhouette width, it is not the best algorithm for clustering on the dataset available to us. The underlying meaning of clusters over time and the reasons behind statistical results are discussed and analyzed in detail, with the aim to highlight possible student group structures present in a high school class.

Research Strand

Data Science for Learning Processes and Education

Primary author(s) :   FARNÈ, Matteo (University of Bologna)

Co-author(s) :   Mrs. TARABORRELLI, Gioia

Presenter(s) :   FARNÈ, Matteo (University of Bologna)

Session Classification :   Assessing student competence: current methodological issues in case studies (Invited Symposium)

Track Classification : Assessing student competence: current methodological issues in case studies (Invited Symposium)

Contribution ID : **15**   Type : **Invited talk**

ASSESSING THE GENDER GAPS IN MATHS COMPETENCE: AN OVERVIEW OF WHAT WE KNOWN FROM INVALSI DATA

martedì 31 gennaio 2023 12:30 (45)

Amongst the extensive research conducted in the context of gender gaps in educational context, an issue that is frequently examined is the role of pre-university education and of the study of STEM subjects, in particular mathematics. The impact of how this discipline is taught, along with the relative attitudes to and interest in this subject, seem to be the main factors pushing young people, and girls in particular, away from the study of STEM subjects. In Italy several studies have been conducted to investigate the gender gap in mathematics using standardized large-scale data from INVALSI tests. The presentation will discuss a review on recent learning results in mathematics of young Italian students both in a cross sectional context and in a longitudinal perspective. For cross-sectional data the focusing on the differences in top and low performances may confirm or not whether the gender differences are more evident within groups of students characterize by different achievement level. The cross-section analyses provide only snapshots of the level of competencies acquired at a specific grade in a given year. On the contrary, a longitudinal approach could evaluate how the performances change over time. Besides, INVALSI makes data available including not only the test responses, but also a set of variables dealing with socio-demographic and economic characteristics, the educational path and, for some grades, student self-reported information about individual and emotional aspects. In the literature, an interesting issue deals with the emotional aspects induced by the test administration. The emotional component behaviour can partially explain the gender gap performance. The results of the illustrated studies offer room to enrich the debate around the international recommendations to promote gender equality in education by ensuring equal opportunities in making educational choices between boys and girls and making the study of STEM subjects equally inclusive and attractive

Research Strand

Primary author(s) : Prof. MIGNANI, Stefania (Department of Statistical Sciences, University of Bologna)

Presenter(s) : Prof. MIGNANI, Stefania (Department of Statistical Sciences, University of Bologna)

Session Classification : Keynote 5

Track Classification : Keynote 7

Contribution ID : **16**   Type : **Oral Contribution**

Do the epistemological bases of qualitative and quantitative paradigms in STEM education still hold today?

lunedì 30 gennaio 2023 11:20 (20)

Recently, the Physical Review Physics Education Research journal organised two calls for papers focused collection to "examine and challenge" quantitative (2017) and qualitative methods (2021) in that field. Both calls had a subtitle, "a critical examination", which foretells the need to focus on these two methodological paradigms, not taking them for granted, inviting to reposition them in today's panorama.

The division of research into these two methodological paradigms originates in the philosophical debate, occurring at the end of the nineteenth century, which aimed to confront the nature of humanistic knowledge with the scientific one. Despite its epistemological origin, the debate between qualitative and quantitative approaches within the research in science and STEM education focuses mainly on pure methodological aspects:

-What methodological approaches are best suited to answer certain research questions?

-Which methods could be used with certain types of data?

-What methodological approaches could best guarantee specific characteristics of the results (e.g. explaining a cause-effect relationship, being objective, specific, general, etc.)?

The aims of the presentation are: to point out the epistemological assumptions that stay behind the traditional qualitative and quantitative methodological approaches; to argue why the big data revolution has been questioning the basic epistemological assumptions that justified the distinction in the past; to raise examples of foundational epistemological questions that can motivate the need to overcome the traditional dichotomy and ground the elaboration of new "mixed" methods.

The first aim is pursued by presenting the results of the historical debate on the methods in terms of its ontological and epistemological pillars: the view of reality that justified the distinction between natural and human realms; the "normative or idiographic" nature of the laws studying a phenomenon. Then, we will show how the different possible ontological and epistemological positions influenced the methodological principles, chosen by the different traditions to make knowledge reliable, general, robust, effective, objective and "true".

The second aim is pursued by discussing some examples taken from research in the way of science education of how sometimes this dualism doesn't seem to hold, or at least doesn't differentiate to make sense. The advent of online learning platforms and extensive use of technologies in classrooms (e.g. sensors) leads to big data about students' behaviour being available in educational research. These particular cases raise epistemological and ontological questions about which reality we are looking at and how we could stand toward this reality (made of real and virtual ones). Different ontological attitudes are emerging, for example in the EDM, and LA communities, and in Informed ML. Moreover, the use of big data, extended and rapid by nature, in the context of STEM education research, which is specific and detailed, makes qualitative or quantitative approaches not completely suitable.

Finally, questions are raised to stimulate reflections to re-think about the epistemological approaches to methods in the field of STEM education such as "what new methodological paradigms are needed in order to deal with these research changes and meaningfully characterise them?" and "Can we still talk about natural settings in modern learning environments?".

Research Strand

Methods and Methodological Aspects in Science Education Research

Primary author(s) :   CARAMASCHI, Martina (Department of Physics and Astronomy - University of Bologna)

Co-author(s) : Dr. FANTINI, Paola (Department of Physics and Astronomy - University of Bologna); Prof. LEVRINI, Olivia (Department of Physics and Astronomy - University of Bologna)

Presenter(s) :   CARAMASCHI, Martina (Department of Physics and Astronomy - University of Bologna)

Session Classification :   Issues in Assessing STEM and Math Education (Contributed Symposium)

Contribution ID : **17**   Type : **Invited talk**

Exploring the impact of student characteristics and social context on mathematical literacy

lunedì 30 gennaio 2023 14:15 (30)

The diagnosis of the a-priori weaknesses of students based on their features and past experience (e.g. age, sex, type of diploma) together with the features of the social context represents one of the preliminary assessment activities that can be taken into consideration so as to define different typologies of students and then develop a tailored learning programme for each student group. This paper is based on the study of a specific aspect of the school-leaving student's ability, namely mathematical skills. It is a matter of fact that the results of several assessment exercises show that the mathematical knowledge of students leaving secondary school is becoming worse and worse. The aim of the paper is to evaluate if and how the socio-demographic features of students impact on their mathematical ability. Exploiting data of the INVALSI tests and the potentialities of quantile regression we will also explore if this impact is different in case of observed or unobserved heterogeneity of the student population.

Research Strand

Primary author(s) :   DAVINO, Cristina (Dipartimento di Scienze economiche e statistiche -)

Co-author(s) :   Prof. VISTOCCO , D.; Prof. ROMANO, Rosaria

Presenter(s) :   DAVINO, Cristina (Dipartimento di Scienze economiche e statistiche -)

Session Classification :   Multilevel Analysis for research on STEM Education (Invited Symposium)

Contribution ID : **18**
Type : **Invited talk**

An analysis of the differences in Italian students' performance in STEM and no-STEM courses

lunedì 30 gennaio 2023 14:45 (30)

In recent years, exploring the determinants that may influence students' achievement has received much attention. Empirical studies have found that the most important factors which affect student performance are students' characteristics, family background, school attended, and regional residence. This paper aims at investigating the differences in students' performances among Science, Technology, Engineering and Mathematics (STEM) courses by using regression models. To measure performance (university success), we focus on the number of ECTS credits earned during the first year, since it represents an important moment in the students' path at university. The analysis concerns students enrolled at 3-year STEM degrees in an Italian university located in the South of Italy during the last 5 years, with a focus on the number of university credits earned during the first year (a good predictor of the regularity of the career). In particular, the main purpose is to estimate the probability of getting at least a certain number of credits at the end the first year and identify the factors which might affect it. As threshold, we opt for choosing that fixed by the Italian National Agency for the Evaluation of Universities and Research Institutes (ANVUR) in the Annual Review Report ("Scheda di monitoraggio annuale") for evaluating the university courses. The data are collected from the Student Information System (ESSE3), a student management system used by most Italian universities, which manages the entire career of students from the enrollment to graduation. It contains information about students' high school diplomas, personal characteristics, exams, abroad experience, internship, and degrees. Using students' identification numbers, we merged the dataset which contains students' demographic information and the dataset which reports all exams taken by students. Since the aim is to estimate the probability of earning a given number of credits, the dependent variable is the number of credits accrued at the end of the first academic year, while the independent variables, among the others, are the gender, course of study, type of diploma, final high school mark, students' age, residence place, average exams mark, and the distance between residence place and university.

Research Strand

Primary author(s) : RESTAINO, Marialuisa (DISES Università di Salerno)

Co-author(s) : LA ROCCA, M.; NIGLIO, M.

Presenter(s) : RESTAINO, Marialuisa (DISES Università di Salerno)

Session Classification : Multilevel Analysis for research on STEM Education (Invited Symposium)

Contribution ID : **19**   Type : **Invited talk**

"I don't know": dealing with uncertainty in students' perceived competence in STEM

lunedì 30 gennaio 2023 15:15 (30)

The importance of STEM (science, technology, engineering, and mathematics) subjects has been growing in different economic sectors and in general for society as a whole (Murphy et al., 2019). Nevertheless, there are significant concerns in STEM education (Höhne et al., 2019) mainly related to the students' difficulties in achieving successful academic performances, the high dropout rates, and a consistent gender discrepancy (Priulla et. al., 2021). Students' motivation to enrol in STEM courses is generally related to career expectation and their perceived competence in scientific subjects (Franks et al., 2019). Some students are, however, cautious about or unsure of their abilities to succeed in mathematics and science, leading to "don't know" responses when questioned about their competencies. The aim of this contribution is to analyse uncertainty in students' perceived competence from an integrated assessment perspective. We employ a within-item multidimensional latent class IRT model (Bacci et al., 2014) to analyse item responses about self-concept in mathematics and science, taking into account the "Don't Know" option.

Research Strand

Primary author(s) : FABBRICATORE, Rosa (Università degli Studi di Napoli)

Co-author(s) : Dr. SANTELLI, F.; Dr. PALAZZO, L.

Presenter(s) : FABBRICATORE, Rosa (Università degli Studi di Napoli)

Session Classification : Multilevel Analysis for research on STEM Education (Invited Symposium)

Contribution ID : **22**   Type : **Oral Contribution**

Investigating secondary students' identification with physics through structural equation modeling

lunedì 30 gennaio 2023 16:55 (20)

In STEM education, the identity framework is often used to investigate students' intention to pursue a STEM-related career. The purpose of this study was to explore the relationships between physics identity, interest, recognition, and performance-competence in physics. The analysis was based on a Likert-scale survey aimed at measuring the addressed constructs, administered online to N = 1135 Italian undergraduate and high school students. We validated a structural model in which performance-competence play the role of the independent variable, physics identity represents the dependent variable, while interest and recognition act as total mediators. We also considered two moderating variables for this model, gender and previous experience with physics, finding significant gender differences only for the path mediated by recognition, while very significant differences were found depending on different experience with physics. Results have implications for instruction in terms of understanding the mechanism underlying the promotion of students' identity development in physics.

Research Strand

Primary author(s) : Dr. CATENA, Danilo (University of Udine)

Co-author(s) : Prof. TESTA, Italo (Department of Physics "E. Pancini")

Presenter(s) : Dr. CATENA, Danilo (University of Udine)

Session Classification : Measurement of cognitive and affective variables in STEM (Parallel Contributed Symposium)

Track Classification : Session 2a – Measurement of cognitive and affective variables in STEM

Contribution ID : **23**   Type : **Oral Contribution**

Use of Multiple correspondence analysis to investigate students' mental models about quantum entities

lunedì 30 gennaio 2023 17:15 (20)

Research Strand

Primary author(s) :   GIULIANA , Giovanni (University of Camerino)

Presenter(s) :   GIULIANA , Giovanni (University of Camerino)

Session Classification :   Measurement of cognitive and affective variables in STEM (Parallel Contributed Symposium)

Track Classification :   Session 2a – Measurement of cognitive and affective variables in STEM

Contribution ID : **32**   Type : **Invited talk**

A Study Of Gender Differences In The Selection Of Stem Courses At University With Large Scale Assessment Data Of Invalsi

martedì 31 gennaio 2023 14:15 (25)

In the most recent years an effort of INVALSI has been addressed to the match of its large scale population data with other sources of administrative/survey data. Technical difficulties and privacy limitations made this process complex and challenging, however, a first linked dataset has been recently developed joining together data from INVALSI, MIUR and University Register . This joint dataset allows researchers to have a wide set of panel data in which each student is followed from primary school until university enrolment and final exam. About 248.000 university students enrolled in 2019/2020 are retrospectively match with their INVALSI test in the previous year (2018/19) and with Ministry of Education Dataset in the same year, with a rate of perfect match of 96%. The whole collection of data in the joint dataset included: • basic anagraphic data, such as gender, migration background, socio economic status as well as parental education (INVALSI dataset, validated through MIUR data on the same students) • Performance at cognitive tests (INVALSI dataset) • Characteristics of the school career and performance of the students: typology of high school, regularity, teachers' marks (INVALSI dataset) • Information about university career: year of enrolment, class of discipline, date and mark of final examination In this work we present preliminary research lines using the joint dataset to gain a better understanding of gender inequalities in the choice of STEM at university. The reference population is a subgroup composed only by students attending a scientific lyceum in grade 13 of 2018/19, who are supposed to pursue with a scientific discipline at university. The study delineates the significant differences between girls and boys in the selection of STEM (42,6% versus 54,6%), with a difference of about 20% in favour of boys. The most analytical part of the research deals with models that aim to evaluate relevant factors for choosing STEM at University, separately for girls and boys. We built up logistic regression models, stratified for gender, and estimated the impact of factors pertaining to the demographic, socio-economic and school/skills domains. In the development of these models, we address and discuss some methodological issues, about multicollinearity of socio-economic variables and selection of best model with AIC criterion. Our research highlights that cultural factors, i.e. mother's educational attainment have a stronger positive impact in the choice of STEM compared to socio-economic factors. In a similar fashion, in the selection of a highly scientific course, the mark in Mathematics given by teachers is more influential than standardized test scores obtained in the INVALSI test; with a predicted probability of choosing STEM increasing from 0.30 to 0.62 passing from weak to excellent mark in Mathematics.

Research Strand

Primary author(s) : FALZETTI, Patrizia (INVALSI)

Co-author(s) : Dr. GIANNANTONI, Patrizia

Presenter(s) :   FALZETTI, Patrizia (INVALSI); Dr. GIANNANTONI, Patrizia

Session Classification :   Use of INVALSI data for research on STEM Education (Invited Symposium)

Contribution ID : **33**   Type : **Invited talk**

Open Research Using Invalsi Data And Multilevel Models

martedì 31 gennaio 2023 14:40 (25)

The purpose of the session is to show how, thanks to the opportunities disclosed by the development of practices related to the idea of Open Research, it is possible (even for an independent researcher) to address issues of great cognitive and practical importance through the combination of high-quality open data, accessible computing tools and advanced statistical models. Specifically, the data taken into consideration are INVALSI open data on students' mathematic skills; the data manipulation resource is R, a high-level programming language whose interface can be freely downloaded and installed by everybody; the statistical tools belong to the family of multilevel regression models, which can now be implemented through very sophisticated Bayesian algorithms. The combined use of data, resources and models will be illustrated and developed in connection with the treatment of a very delicate problem: the possibility that, in our country, students with a migrant background may suffer a penalty which is not justified by their actual level of skills and abilities but, rather, depends on some sort of bias rooted in the evaluation systems implemented by the schools. The topic is of particular interest since the share of foreign students (or students with a migrant background) is expected to increase sharply in the next decade. To give adequate relevance to the issue, it must be observed that in our Country schools are the main – some would say the only – agency for the integration of immigrants. Therefore, the perspectives of our society crucially depend on how efficiently schools will be able to carry out their tasks. The INVALSI data allow us to approach the problem of schools' fairness in an objective way. The standardized nature of INVALSI tests helps us to try to identify the presence of potential bias in the overall evaluation system. A possible research question can be formulated in these terms: Invalsi scores being equal, do Italian and foreign students receive (approximately) the same marks by their teachers? The INVALSI score is aimed at assessing fundamental abilities – for example general mathematic skills – thorough the administration of standardized tests; on the other hand, marks are the outcome of a complex evaluation carried out by teachers – a kind of assessment that considers many other factors, for example behavioural ones. Any discrepancy between standardized scores and marks could be a sign of limitations in the former, which is plausible, but also a clue of some bias in the way schools evaluate students with a migratory background (which cannot be excluded on an a priori base: after all, despite a growing number of foreign students almost 100% of our teaching body is made by Italian teachers). In order to address these issues, and provide a contribution to the discussion, it is necessary to approach the data with tools that allow us to extract as much information as possible from them, without introducing further distorting factors. This is essentially an open process, which must be subjected to strict critical supervision and requires transparency; the tools and techniques that will be illustrated seek, precisely, to build a transparent path to open discussion and criticism.

Research Strand

Primary author(s) : LORENZO , Maraviglia (INVALSI)

Presenter(s) : LORENZO , Maraviglia (INVALSI)

Session Classification : Use of INVALSI data for research on STEM Education (Invited Symposium)

Contribution ID : **34**   Type : **Invited talk**

A Counterfactory Approach To The Evaluation Of Educational Policies: Methodology, Variables And Results

martedì 31 gennaio 2023 15:05 (25)

The fight against early school leaving, the prevention of school failure, the raising of basic skills levels and support for learning are among the main objectives of the interventions financed through projects activated in favor of schools. The study is aimed at quantifying the effect on the levels of competence achieved by students in the INVALSI tests through interventions aimed at (i) improving school performance and (ii) reducing early school leaving through the construction of a longitudinally anchored and counterfactual analytical system able to evaluate the progression and/or changes in students' skills over time in three areas: Italian, Mathematics and English. This evaluation is carried out in a comparative way between the students of the schools involved in the Projects (hereinafter "treated cases") with respect to those not involved (hereinafter referred to as "untreated cases"). The purpose of effect evaluation is to verify the ability of an intervention to change the behavior or conditions of a specific target population in the desired direction. The method used, the Difference-in-Difference (Angrist & Pischke, 2009; Keele, 2020), makes it possible to estimate the value of the outcome variable of the students treated in the event that they would not have received the treatment (counterfactual). The net effect of the treatment is therefore obtained as the difference between the score observed on the "treated" students and the score that would have been observed in the absence of the treatment. To estimate the effect of interventions on the results of standardized tests, a quasi-experimental evaluation design was introduced by adopting a counterfactual approach. In counterfactual analysis, three groups are referred to: (i) control group, (ii) factual group, and (iii) counterfactual group. The control group, built using the Propensity Score Matching technique (Martini A., 2011), is made up of students who have not received the treatment and who present characteristics (socio-demographic, such as gender, age, geographical location, citizenship, socio-economic status and culture of the family of origin; and scholastic, previous level of competence and regularity with respect to the course of studies) completely similar to the characteristics of the students who received the treatment and who make up the current group. Finally, the counterfactual group, constructed with the help of a control group, is the one for which it is estimated what the variation over time of the students' competence in the treated group would have been if they had not received the treatment. The data used selected from the national surveys to which INVALSI administers to students in the various school levels covered by the Survey, thanks to the presence of a longitudinal student code, it was possible to follow the students over time to evaluate the variation in the outcomes following specific interventions. It was also possible to proceed with a clustering of the group of treaties to evaluate the effectiveness of policy interventions on sub-populations on the basis of the different characteristics present within them. Thanks to the subdivision into subpopulations, it was possible to conclude that the efficacy of the treatment had different outcomes according to the identified subgroup. The results will also be discussed based on the literature on the subject.

Research Strand

Primary author(s) : BENDINELLI, Andrea (INVALSI)

Presenter(s) : BENDINELLI, Andrea (INVALSI)

Session Classification : Use of INVALSI data for research on STEM Education (Invited Symposium)

Contribution ID : **35**   Type : **Invited talk**

The Calculation Of Socio-Economic-Cultural Status Indicator Of Learners From Invalsi Data

martedì 31 gennaio 2023 15:30 (25)

The socio-cultural and economic characteristics of students and the families from which they come play a very important role on the learning levels achieved, even from the early years of schooling. It is well established, in fact, that students who live in conditions of higher economic as well as social and cultural advantage have a better chance of achieving more satisfactory results during their education. It is now well known in the literature that the socio-cultural and economic condition, the so called background, has a significant predictive value on learners' achievement (INVALSI, 2008; OECD 2007). Having established this, it is essential to have an instrument, i.e., an indicator, to measure in some way the background of the learners of interest. The definition of an indicator of socio-economic-cultural status and its calculation poses general and technical problems. This paper presents the method and techniques used to calculate an indicator of socio-economic-cultural status (ESCS) of pupils who participated in the National Assessment Service (NES) surveys. The calculation of ESCS is based on discrete indicators such as parents' level of education (HISEI) and their employment status (PARED), but also on a continuous indicator that can express a proximity measure of the material conditions in which the pupil lives outside school (HOMEPOS). The latter indicator is calculated from the data of the regionally stratified sample of students participating in the INVALSI tests and using techniques related to the methodological framework of Rasch Analysis. The paper also illustrates some operational solutions to overcome, at least in part, the problem of missing data, always present in large-scale surveys, especially for so-called context variables. Finally, some initial analyses are proposed about the relationship between ESCS and learning levels achieved in Italian and Mathematics.

Research Strand

Primary author(s) :   CAMPODIFIORI, Emiliano (INVALSI)

Presenter(s) :   CAMPODIFIORI, Emiliano (INVALSI)

Session Classification :   Use of INVALSI data for research on STEM Education (Invited Symposium)

Contribution ID : **36**   Type : **Invited talk**

What I wish I had known as a Rasch beginner....

martedì 31 gennaio 2023 16:00 (45)

This talk will present a summary of what I wish I had known 30+ years ago as I began my Ph.D. studies with Ben Wright and first learned Rasch methods. In this talk I will present key errors beginners commonly make when they start to use Rasch. Naturally I will briefly discuss the ins and outs of such beginner Rasch errors...my goal..to aid Rasch novices so that they can confidently, quickly and correctly use Rasch. I will also share and discuss what my family calls "Bill's Rasch elevator speech"...such a speech is the one I can give very quickly (60 seconds) in an effort to help others understand why Rasch is important. Quickly explaining Rasch to those who do indeed need your Rasch help (but they do not know they need your assistance) is a critical skill that one must develop. It is an art to succinctly explain Rasch to others, be it in person, or in papers. Even as a beginner with Rasch, by mastering your own elevator speech you can increase the number of people who seek your Rasch expertise, be it in academia or the private sector.

Research Strand

Primary author(s) : Prof. BOONE, William (Miami University)

Presenter(s) : Prof. BOONE, William (Miami University)

Session Classification : Keynote 6

Contribution ID : **37**   Type : **Invited talk**

Evolution acceptance and high school students. Methodological considerations and new perspectives of investigation

lunedì 30 gennaio 2023 09:45 (45)

The idea of biological evolution is not accepted by many people around the world, with a large disparity amongst countries. Some factors may act as obstacles to the acceptance of evolution, such as religion, a lack of openness to experience, and not understanding the nature of science. Although the strength of the association between evolution acceptance and non-scientific factors varies among studies, it is often assumed that resistance to evolution is the by-product of a religious background. Some studies are even more specific and try to associate the acceptance of evolution with precise religious affiliations. In my speech I will propose an introduction of the main tests used internationally to measure evolution acceptance and some statistical tools that allow our research team to verify the relevance of sociocultural factors in predicting such acceptance. Starting from an in-depth reflection on the factors that influence the knowledge and acceptance of evolutionary theories, the Italian-Brazilian research group formulated a research question to address this complex picture considering the same religious affiliation in two different countries with deep sociocultural differences. Catholic Christians in Italy and Brazil have several similarities, including many family connections owing to immigration history. Brazil is the country with the highest number of Catholic Christians in the world, and Italy is the hub of Roman Catholicism. These conditions allowed to conduct a survey on the adolescent population with two statistically significant national samples in 2014. We adopted a clear definition of evolution acceptance despite the complexity of the discussion on the subject in different languages, taking acceptance as the expression of explicit recognition of the objective validity of known scientific statements about evolution under absolute anonymity. This definition considers two steps. The first is associated with scientific statements about evolution, which must be clear and well known, avoiding issues under discussion, for instance, about the origin of life. Students must show not only a positive attitude towards evolution but also express clearly that a statement based on biological evolution is considered a valid premise to construct a judgement about the real world. The second step refers to objective conditions in which a person may admit his/her positive judgment about a certain scientific statement. We aimed to explore the strength of associations among nationality, religion, and the acceptance of evolution by students using multiple correspondence analysis (MCA) and statistical tools, with nationwide samples from two different countries. In our research we found that wider sociocultural factors predict the acceptance of evolution to a higher degree than a religious background. Roman Catholic students showed significant differences between the two countries, and the gap between them was wider than between Catholics and non-Catholic Christians within Brazil. Our conclusions support those who argue that religious affiliation is not the main factor in predicting the level of evolution acceptance. The sociocultural environment and the level of evolutionary knowledge seem to be more important in this regard. These results open up new interpretative perspectives and provide a better understanding of attitudes towards evolution.

Research Strand

Presenter(s) :   Prof. PELLEGRINI, Giuseppe (Observa Science in Society)

Session Classification :   Keynote 1

Contribution ID : **38**   Type : **Invited talk**

Clustering methods to study student reasoning lines: theoretical aspects and experimental results

lunedì 30 gennaio 2023 12:30 (45)

Studying reasoning lines students deploy when dealing with problematic situations is, in Education Research, one of the most important aims. The reasoning lines can be achieved by analysing the answers that students give to a questionnaire, although it becomes increasingly complicated as the number of students to be analysed increases. In my speech I focus on a quantitative method based on clustering by discussing theoretical and methodological aspects. The method can allow a researcher to analyse a set of answers given to a questionnaire, even in the case of a large sample of students. Moreover, clustering is not presently common to the physics education research community. I describe in detail two different clustering methods, a hierarchical one and a non-hierarchical one. I introduce a binary coding that makes the answers quantitatively analysable. Moreover, a correlation coefficient and a metric suitable for measuring student similarity in the case of binary coding are presented. Then, criteria for choosing the optimal number of clusters for both the clustering methods are discussed. For the same purpose, a new coefficient is introduced to measure the total amount of information one can obtain from a clustering solution. I show that each cluster can be characterised by its centroid. It summarizes the most frequent answers given by students in a given cluster. An example of a clustering procedure for experimental data is given. The comparison between the results obtained through the two clustering methods shows a good agreement exhibiting robustness in the proposed method.

Research Strand

Presenter(s) :   Dr. BATTAGLIA, Onofrio Rosario (Università di Palermo)

Session Classification :   Keynote 2

Contribution ID : 39　　　　　　　　　　　　　　　　　Type : **Invited talk**

Assessment of social and emotional skills in cross-national settings: an OECD approach

martedì 31 gennaio 2023 09:00 (45)

Research Strand

Presenter(s) :　Prof. KANKARAS, Milos (Faculty of Philosophy, Department of Psychology, University of Montenegro, Montenegro)

Session Classification :　Keynote 3

Contribution ID : **40**   Type : **Invited talk**

The black box is not that black anymore: The use of eye-tracking in STEM education research

martedì 31 gennaio 2023 09:45 (45)

Despite many attempts to measure pupils and students' learning outcomes, their performance in problem solving has largely remained something of a black box. This talk will therefore focus on the potential of using eye-tracking data in combination with the think-aloud method. Case studies will be used to show how to analyze data from actual science education research. Finally, further possibilities of using eye-tracking in STE(A)M education research will be outlined.

Research Strand

Primary author(s) :   Prof. RUSEK , Martin (Charles University, Faculty of Education. Prague, Czech Republic)

Presenter(s) :   Prof. RUSEK , Martin (Charles University, Faculty of Education. Prague, Czech Republic)

Session Classification :   Keynote 4

Contribution ID : **41**   Type : **Invited talk**

Automated Assessment Of Open-Ended Question Of Invalsi Tests

martedì 31 gennaio 2023 16:45 (45)

This work describes the new procedures of automated corrections of freeform answers given by the 8th, 10th and 13th grade students to open-ended questions in CBT Computer Based Test) INVALSI tests. INVALSI team, composed of statistical and computer scientists, responsible of open-ended question correction, has implemented an algorithm to process text strings of different complexity. Before survey distribution, the correction team and the items authors group discuss to define the correction criteria, that is a set of rules to determine the correct or incorrect classification for each answer given by the students for a specific item. The discussion produced, moreover, the indications to remove useless elements for the classification, then translated in operations of the algorithm on the textual data such as punctuation detection and removal, special characters, articles, conjunctions, word lemmatisation, etc. The answer strings were subsequently processed by a "data cleaning" operation, that was focused on the automated correction of spelling and typing errors, by detection and substitution of "out-of-vocabulary" words (OOV words). After the "data cleaning" phase, the correction criteria fixed by the experts have been translated in logical IT patterns, aiming to uniquely defining the set of admissible ways to give a correct answer. The last test phases of the algorithm were characterized by a constant exchange of information about the encoding, among the authors' team and the correction team, this passage being critical to refine the logical rules used for correction and to get more consistency and precision between the encoding produced by the algorithm and the authors' indications. The final test of the algorithm ends with a comparison between the manual encoding by video correction and the one processed by the algorithm on a set of items already processed in a former test: the algorithm is accounted as accurate enough and aligned to the indications of authors' team when the complete accordance of the two encoding was achieved. The methodological approach, countable as a method of supervised automated correction, represents a valid compromise between a manual encoding and a totally automated one, typical of the machine learning algorithms. This method has indeed the benefit of considerably reduce the hours/man needed to correct the open-ended answer items, when compared to a manual procedure, and get a better accuracy reducing the wrong encoding matches, when compared to a non-supervised automated procedure. A comparison between supervised and non-supervised automated procedure has been eventually done to evaluate the distance between the two methodological approaches.

Research Strand

Presenter(s) : Dr. MARSILI, Michele (Invalsi)

Session Classification : Keynote 7