Article

Making STEM Education Objectives Sustainable through a Tutoring Program

Francisco-Domingo Fernández-Martín [1],*, **José-Luis Arco-Tirado [1]**,
Francisco-Javier Carrillo-Rosúa [2,3], **Mirian Hervás-Torres [1]**, **Juan-Francisco Ruiz-Hidalgo [4]**
and Carmen Romero-López [2]

[1] Department of Developmental and Educational Psychology, University of Granada, 18071 Granada, Spain; jlarco@ugr.es (J.-L.A.-T.); miriamhervas@ugr.es (M.H.-T.)
[2] Department of Didactics of Experimental Sciences, University of Granada, 18071 Granada, Spain; fjcarril@ugr.es (F.-J.C.-R.); romer@ugr.es (C.R.-L.)
[3] Andalusian Institute of Earth Sciences, CSIC-University of Granada, 18100 Armilla, Granada, Spain
[4] Department of Didactics of Mathematics, University of Granada, 18071 Granada, Spain; jfruiz@ugr.es
* Correspondence: fdfernan@ugr.es

check for
updates

Abstract: The objective of this research was two-fold. First, to determine the impact of a Science, Technology, Engineering and Mathematics (STEM) Education program on school performance amongst primary education students. Second, to identify the potential benefits of this program on the key competences of university students in Primary Education Teacher Training. The primary education students' sub-sample, after being matched on key covariates, was randomly assigned either to the experimental ($N = 25$) or control group ($N = 25$). The university students' sub-sample consisted of 26 students self-selected from the Primary Education Teacher Training degree. The intervention consisted of 20 two-hour weekly sessions of highly structured after-school tutoring delivered by previously trained university students. Although statistical significance was not reached for the hypotheses tested, notably, the results showed between small and moderate effect sizes (i.e., magnitude and direction of the program impact) for primary education students on core STEM subjects (e.g., mathematics $d = 0.29$, natural science $d = 0.26$), and for university students on some key employability competencies (e.g., action orientation $d = 0.27$ or team orientation $d = 0.54$).

1. Introduction

In spite of the impact of the economic and financial crisis on labor markets, the demand for employment in sectors linked to Science, Technology, Engineering and Mathematics (STEM) has been increasing significantly in the last decade [1,2]. Furthermore, this growth trend is expected to persist in the coming years, and in Spain a rise of 9.68% is expected in the number of employees that the scientific-technological sector will demand by the year 2025 [3].

However, in the European Union the number of students accessing STEM degrees (i.e., life science, physical science, mathematics and statistic, computing, engineering and engineering trades, and manufacturing and processing) is not increasing, the university graduation rate of scientific-technological degrees remains stable, and the number of STEM vocational training graduates is decreasing [4,5]. More specifically in Spain for example, the number of students enrolled in these scientific-technological university degrees is decreasing drastically, while the students who do enroll in these studies and training programs show a low academic performance [6]. This decrease in the

number of students and graduates in STEM degrees is the result of the combination of several social, cultural, economic, and educational factors. The literature points out that the main causal variables are the following [4,7–9]: students' disinterest in the scientific-technological field during compulsory educational stages, stimulated by their level of self-efficacy in the STEM curricular areas; the fact that they belong to ethnic or cultural minorities; their socio-economic status; and the high rates of failure and drop-out that these degrees present [10].

Nevertheless, the gap between the demand for STEM professionals with a certain level of qualification by the labor markets and the number of students that access and graduate in STEM degrees is causing certain recruitment problems in the scientific-technological sector, where certain jobs are not being filled due to the absence of professionals with the required level of qualification [2,4,5]. This situation has become a priority for a large number of government officials, since it is considered a key element for promoting the level of competitiveness and economic growth of any country [11].

To this end, several political, practical, and research strategies and measures have been implemented, leading toward different initiatives aimed at [4,12–16]: (a) reforming the STEM curriculum, especially in terms of teaching methodology, among which inquiry-based science education (e.g., the INQUIRE, Mind the Gap, SPICE, and PRIMAS projects) is highlighted, as well as the push for greater contextualization in science education (e.g., integrating science education within current social issues); (b) improving the training and professional development of STEM teachers through various means, such as delivering formal continuous professional development activities for teachers by science centers, creating networks for the exchange of good practice in the field of science teaching, the inclusion of key objectives specifically related to improving teacher education in the strategic frameworks in place to promote science education, or the implementation of initiatives that target initial education of science teachers (e.g., the STELLA and GRID projects); (c) improving the vocational and professional guidance that the students receive through initiatives aimed at challenging the negative perception of science (e.g., the MOTIVATION and YOSCIWEB projects), or providing high quality guidance and good labor market information (e.g., career guidance at transition points from one level of education to another); and (d) the need to create sustainable learning environments in which students engage in authentic activities guided by frequent, significant, and meaningful formative assessment through initiatives that aim to develop effective cooperation with employers (e.g., school and business partnerships), or develop skills for forecasting and monitoring systems. For example, in the framework of the European Union, many initiatives can be found, among which Scientix, InGenious, and the STEM Alliance, led by the European Schoolnet, stand out [4,13,14]. However, in Spain there currently is no specific strategy to address this need, although it is true that educational regulations do place special emphasis on the improvement of teaching in the scientific-technological field, and there are certain institutions delivering teaching and training courses (e.g., National Institute of Educational Technologies and Training) [14].

Indeed, the Nurture Thru Nature (NtN) program was developed to serve this specific and important need highlighted by the Sustainable Developmental Goals [10]. From this perspective, several authors [17,18] have developed STEM education as an active, interdisciplinary, and integrating focus of the teaching-learning process of these four major areas of knowledge, which is aimed at solving problems and innovation with the purpose of increasing the interest of young people for STEM studies and professions. This approach is based on the Inquiry-Based Science Education (IBSE) teaching method, which has been recommended by European institutions [19,20], and its use has expanded appreciably in Europe, although there are only a very few Spanish schools using it [21]. Although different typologies and variants of this teaching method have been recognized (e.g., model-based inquiry, argument-driven inquiry or 3D inquiry), and different studies showed evidence of their effectiveness [22], recent meta-analyses demonstrated that guidance is pivotal to successful inquiry-based learning at all educational levels [23]. IBSE typically involves the following [24]: (a) scientifically oriented questions that engage students; (b) students compile evidence, mainly from experimental procedures, to evaluate their explanations to the problems; (c) students formulate

explanations from evidence to scientifically address oriented questions; (d) students evaluate their own previous explanations that frequently correspond to misconception and alternative ideas, reflecting scientific understanding; and (e) students communicate and justify their proposed explanations to the problems.

1.1. The NtN Program at the University of Granada (NtN-Granada)

The NtN-Granada is an adaptation of the NtN-Rutgers program resulting from the inter-institutional research collaboration between Rutgers University (USA) and the University of Granada (Spain). The main characteristics of both programs are collated in Table 1, with NtN-Rutgers showing a high level of effectiveness in four editions, with a considerable increase in the knowledge and school performance of the participating students in the STEM subjects [25–27].

Table 1. Summary of key components of the Nurture thru Nature (NtN)-Rutgers and NtN-Granada programs.

NtN-Rutgers (USA)	NtN-Granada (Spain)
Program	
NtN-Rutgers, Partnership between Rutgers University, New Brunswick Public Schools, and Johnson & Johnson.	NtN-Granada, which is intended to validate the NtN-Rutgers' impact model in the Spanish socio-educational context.
Program type	
After-school and summer; moderate economic and logistic resources; availability of material such as a garden.	After-school; very limited economic and logistic resources; everyday objects frequently used; lack of resources such as gardens.
Aims	
To stimulate the development of human and social capital among compulsory education students in risk of social exclusion.	To stimulate the development of human and social capital among compulsory education students in risk of social exclusion and to improve primary education teacher training.
Implemented by	
Rutgers faculty/staff and students (granted).	University of Granada faculty and students (volunteers).
Methodology and methodological orientations	
- Direct instruction, in the case of school support activities. - Inquiry and hands-on learning for the scientific-technological activities, in which situations of the daily life of the students are approached; "head, heart and hands" model [26], a garden and outdoor-based curriculum (orientation to environmental and sustainable education perspective) [25,27].	- Direct instruction, in the case of school support activities. - Inquiry and hands-on, oriented, through IBSE methodology and STEM orientation, for scientific-technological activities. - Peer tutoring and service-learning.

NtN-Rutgers and NtN-Granada are inspired by the "Head, Heart and Hands" environmental educational model for transformative learning that informs the current educational movement to foster environmental understanding and sustainability [25–27]. This model is based on Dewey active learning philosophy, especially in two of its principles (i.e., "the need to connect a student's prior knowledge and experience, however limited, to current and future learning experiences, and the need to situate the student's learning in the here and now, providing opportunities to apply mathematics and science knowledge to every situations") [27], and it is designed to promote student learning through the simultaneous involvement of cognitive, emotional and psychomotor domains [26]: "The model shows

the holistic nature of transformative experience and relates the cognitive domain (head) to critical reflection, the affective domain (heart) to relational knowing and the psychomotor domain (hands) to engagement". Moreover, these programs combine "hands-on" science and environment-based curriculum and they include after-school program elements and components [25–27].

One of the main differences between both programs is that the NtN-Granada program additionally aims to benefit teacher training improvement through incorporating volunteer university students as tutors to deliver the program sessions, rather than the hired staff of NtN-Rutgers. There is a need to provide university students with access to high-quality learning and competencies development opportunities, which aligns with the competency-based education model embedded in the higher education area in Europe. Furthermore, this is a way to incorporate service-learning (S-L) and community engagement, as the new wave of outreach programs for universities deliver their third mission [28]. Both methodological approaches were blended into the NtN-Granada intervention design in the following way: (a) S-L, which integrates knowledge and the curriculum's subjects with learning and doing useful and supportive work in the community [29]; and (b) peer tutoring, in which students of different ages and/or with more knowledge and/or skills, after a process of training and practice, through a framework of asymmetrical relationship externally planned by professionals, help and support other students with less knowledge and/or skills for preventive and/or instructional purposes [30]. Amongst the multiple benefits of both ways of understanding and practicing education, the following are frequently cited [29–36]: an improvement in the school performance of participants, their skills and social relationships, an increase in health indicators and professional skills, and improvements in professional skills and the satisfaction of the educators involved. There is also evidence of the benefits that this type of program exerts on the social climate of institutions and communities by increasing the organizational structure and social participation of their members (e.g., increasing the number and quality of interactions between its members and its immediate environment, improving members' well-being and commitment to their work or duties, decreasing leave days and reducing disciplinary issues) [29,34,35,37].

1.2. Hypotheses

The aims of this study were, on the one hand, to determine the impact of a STEM education program on school performance amongst primary education students, and, on the other hand, to identify the potential benefits of this program on the key competences of university students. To these ends, the following hypotheses were established: (1) as a result of the program, there will be statistically significant differences in the average grade of school achievement obtained by the primary education students of the experimental group in relation to those of the control group in the instrumental and STEM curricular areas at the end of the school year; and (2) the university students will improve in statistically significant terms their average direct scores obtained in personality and key competences for employment in post-tests with respect to pre-tests as a result of their participation in the NtN-Granada.

2. Materials and Methods

2.1. Participants

The sample consisted of 76 participants who were divided into two sub-samples, 50 primary education students (NtN tutees), and 26 university students who participated as student-tutors (NtN tutors). The NtN tutees sub-sample was divided into two groups: the experimental group consisted of 25 students, including 11 girls and 14 boys, with an average age of 10.52 years (age range: 10–11 years), and a distribution by educational level of 14 students in the fifth year and 11 in the sixth year. Likewise, the control group consisted of 25 students, with the equivalent distribution by educational level and sex, and with the same mean and age range as the experimental group.

The sub-sample of NtN tutors was made up of 26 students enrolled in the primary education teacher training degree: 12 students in the first year and 14 in the second year. Their average age was

19.58 years, with a range between 17 and 39 years, while their distribution by gender was 20 women and six men.

The sampling selection was based on a non-probabilistic sampling technique [38] and involved four sampling stages. In the first stage, two educational centers and a degree program were chosen as the field of the study due to various reasons: (a) the limited financial, material and human resources available; (b) the program had to be implemented with a sample of educational and socio-demographic characteristics similar to those of the American counterpart (i.e., fifth and sixth grade in a poor urban school), whilst the degree had to be one in which future primary education teachers were being trained; (c) the applied nature of the program as well as the objectives of the intervention were aligned to the center's expectations and needs; and (d) availability and geographical proximity.

In relation to the second stage, once the required institutional permission was obtained, the divulgation and recruitment plan was carried out, including: (a) six group information and coordination sessions with the institutional managers and teachers of the participants schools, in which they were informed about the program and their input, as well as their participation requested in the dissemination and recruitment of primary education students and their families; (b) four group information sessions with the fifth and sixth grade primary education students of the participant schools, in which they were informed about the characteristics of the program including a demonstration of a brief scientific-technological activity, as well as providing them with the family contract agreement (i.e., document similar to a behavioral contract in which the families' rights and obligations were specified, in addition to the authorization from the school to process the child registration in the program) [33]; and (c) two group sessions with the university students from the selected degree, in which they were informed about the characteristics, conditions, costs, and benefits of the program (e.g., voluntary participation, although at the end of the program tutors were awarded with the equivalent of six credit hours by their faculty).

With regard to the third stage, 62 primary education students from the fifth and sixth grades were enrolled in the program by the institutional heads of the participants' schools, after (a) reading and accepting the clauses of the contract agreement of the schools (i.e., document similar to a behavioral contract in which the commitment, rights, and obligations of the schools were specified), and (b) the families were given the family contract agreement, an original copy of the primary education students' report card and the access protocol (i.e., self-report in which the teachers provided the demographic, school and interest information of the students who were enrolled on the program) [33]. Likewise, 29 university students enrolled voluntarily in the program after (a) reading and accepting the clauses established in the NtN tutors agreement (i.e., a document similar to a behavioral contract in which the rights and obligations of university students who enrolled on the program were specified), (b) providing an original copy of their academic report, and (c) completing the NtN tutors protocol (i.e., self-report aimed at obtaining the demographic, academic, and interest information of the university students who enrolled in the program) [30,32], and the Business-focused Inventory of Personality (BIP) (i.e., self-report to measure personality and key competences for employment) [39].

Finally, in the fourth stage, out of the 62 primary education students who initially volunteered, a total of 26 pairs ($N = 52$) were created. Pairs were sorted and matched according to two sets of control variables [40], usually associated with school performance [41,42]: (a) school variables, such as educational center, course, group, specific educational support needs, measures of attention to diversity applied or in progress, and performance in the previous school year; and (b) demographic variables, such as age, gender, nationality, years of residence in Spain, and language used at home. The remaining 10 primary education students were discarded due to the lack of an appropriate match, whilst the members of one of the pairs reported their withdrawal from the program due to problems with their schedules.

In the case of the 29 university students who initially volunteered, 26 were selected as NtN tutors according to the following criteria: (a) a grade point average of more than two points (in a scale of four points), (b) a centile score higher than 50 in the BIP [39], (c) having demonstrated interest and being

available at the requested time intervals, and (d) attending the four training sessions and passing the practical NtN tutor training tests.

2.2. Materials

The instruments used for data collection were the academic report and BIP [39]. An original copy of the primary education students' report card provided by institutional managers of the participating schools was used to calculate school achievement obtained by the primary education students in the instrumental and STEM curricular areas at the end of the school year.

Personality and key competencies for employment was measured by BIP [39], an inventory composed of 210 items with six response levels grouped into 14 scales. It was used because it is one of the tests commonly used to evaluate competencies in training contexts due to its acceptable level of reliability (Cronbach's alpha of between 0.51 and 0.84 points in the different scales) and validity (factorial analysis that demonstrates the factorial structure) in its original version and Spanish adaptations.

2.3. Design and Procedure

The methodological designs adopted were an experimental design of randomized blocks with concomitant variables for hypothesis 1, and a quasi-experimental pre-test/post-test design for hypothesis 2 [40].

Once the NtN tutees' pairs were established, each member of each pair was randomly assigned to either the experimental group or the control group through the online tool called "research randomizer" (https://www.randomizer.org). It was then confirmed that the experimental and control groups were equivalent to the pre-established control variables (i.e., some of them showing the same value and others showing the same proportion in both groups, whereas for the rest of the control variables relevant parametric and nonparametric contrasts were performed, without appreciating statistically significant differences for each of them between both groups). Then, the researchers contacted each family of the control group members to inform them that they had not been selected to participate in the program, although they would remain on the waiting list. Therefore, they did not know their control condition and they did not participate in any of the NtN activities or equivalent (i.e., tutoring sessions: school support and scientific-technological activities). This is a standard decision and protocol followed in this type of program.

The intervention plan was composed of the contents, procedures, and activities extracted from the specialized bibliography [4,12,17,29–35,37,43–45]. In this vein, the NtN tutors' training course (i.e., the first activity of the intervention plan) was aimed at enhancing the university students' skills to efficiently perform the role and functions initially assigned. Thus, four three-hour sessions were delivered in two weeks (2 to 11 November 2016) using an active and participatory methodology to cover the following contents [31,33]: (a) session 1: introduction of the program staff, participants, and training plan, justification of the program, and description and use of the workbooks (i.e., study protocol: a set of materials in which each of the tutoring sessions was presented in a structured manner), (b) session 2: performance of the NtN tutees' needs assessment (i.e., instructions on how to analyze and to review the agenda and school material to establish and prioritize homework, doubts, problems, or difficulties on the instrumental curricular areas), and the design, planning, and execution of school supplemental activities, (c) session 3: the design, planning and execution of scientific-technological activities (i.e., presentation and curricular justification of the activities, and training in the IBSE teaching method: the procedure to be followed by NtN tutors when implementing the activities), and (d) session 4: evaluation and registration of the tutoring sessions, and analysis of potential difficulties or unpredicted situations during the tutoring sessions.

Once the NtN tutors' sub-sample had been selected, they were paired up ($N = 13$) taking into account their matches on the degree program group, program year, academic and personal similarity, and time availability. Next, the NtN tutees of the experimental group ($N = 25$) were assigned to NtN pairs of tutors based on the compatibility of their respective schedules (i.e., two pairs of tutors were

assigned three students each, eight pairs two students each and three pairs one student each), to then start the implementation of the tutoring sessions.

The second activity of the intervention plan was to deliver the tutoring sessions from 14 November 2016 to 11 May 2017, once a week from 16:00 to 18:00 h in the school classrooms assigned by each school. During that time, the first hour was dedicated to the school support activities in one-to-one sessions based on the students' needs (i.e., the establishment of optimal conditions for the development of the session and needs assessment, objectives and planning of the session, providing information illustrated with examples and modeling, supervision of understanding and guided discovery, and independent practice) in the instrumental curricular areas, followed by a 10 min break in which NtN tutees of the experimental group had a snack. Then, the 17 scientific-technological activities previously designed by university staff for this project focused on some contents of the STEM curricular areas (i.e., digestion, musculoskeletal system, fish anatomy, earth materials, cells, microscope function, operation of the microscope, circulatory system, plant biology, pressure and atmosphere, microbes, arthropods, graphics and movements, gravity, types of mixtures and their separation, density and buoyancy, respiratory system and health, volume and capacity, sound and light, electricity and magnetism), following the IBSE teaching method (i.e., identification or establishment of the problem/need/objective, hypothesis formulation, experimentation and information gathering, discussion on the data, and establishment of conclusions on examples of their daily life situations and experiences); this was delivered in small groups (i.e., 2–5 students). In addition, machines and construction of apparatus and basic mathematical operations were transversal contents amongst the scientific-technological activities, which aimed to develop both manipulative and cognitive inquiry skills.

These activities were designed and sequenced in the workbook, taking as a reference the curricular content scheduled during the school year, because they were meant to apply and integrate the STEM curricular contents that the students had learned during their morning classes. In fact, in order to facilitate and align the work of the NtN tutors with that of the teachers in the regular classroom they all had a digital copy of the textbooks. The scientific-technological activities were held weekly in the order established in the workbook, although three of them were implemented in two sessions. Nevertheless, the NtN tutors and the university staff had to introduce some variations in the activities planned to maintain the interest and learning of the NtN tutees of the experimental group, as the specialized literature recommends [30,31,34,35]. Additionally, in order to maximize the impact of the scientific-technological activities delivered, a digital copy of the material with the curricular justification of that activity and the material resources necessary for its implementation and development was prepared (i.e., presentation and initial instructions, concepts to be addressed, and sequence of action) (https://ntngranada.wordpress.com/programa-de-actividades).

Eventually, 20 tutoring sessions were carried out, although some sessions were missed by NtN tutees in the experimental group ($M = 0.96$, $SD = 1.10$) and NtN tutors ($M = 0.85$, $SD = 0.97$).

In parallel to the intervention plan, the monitoring plan to guarantee the fidelity of the program was implemented [31]: (a) a participant observer was included in the NtN tutors training sessions, (b) the presence of program staff in each of the tutoring sessions, (c) three follow-up group sessions between the program staff and the NtN tutors (i.e., analysis of the actions taken in the corresponding tutoring sessions based on the comments of the NtN tutors and the notes registered in their workbooks, providing them with corrective feedback if necessary and improving planning for the next sessions), and (d) a final group session devoted to assessing the overall participation of the NtN tutors in the program, with special emphasis on the difficulties faced and challenges for future editions. Finally, with the evaluation plan of the results, measurements were taken of the dependent variables before and/or after the application of the program to test for statistically and/or educationally significant effects [30,31].

2.4. Statistical Analysis

The data for the different hypotheses were analyzed through the *U* of Mann–Whitney (i.e., a nonparametric statistical hypothesis test to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed) (hypothesis 1) or the *t*-students dependent test (i.e., a parametric statistical hypothesis test to compare differences between one group of units that has been tested twice when the dependent variable is either ordinal or continuous and normally distributed) (hypothesis 2), after applying the Kolmogorov–Smirnov test to check for normal distribution (i.e., a nonparametric statistical test for testing if a variable follows a given distribution in a population). The mean (*M*) and standard deviation (*SD*) were also calculated. The level of statistical significance was expressed as a *p*-value, and we worked with a $p < 0.05$ to establish statistically significant differences. However, the family-wise error rate, resulting from the multiple comparison problems, was controlled with Bonferroni correction (i.e., a multiple-comparison correction used when several dependent or independent statistical tests are being performed simultaneously, and in order to avoid a lot of spurious positives, the alpha value needs to be lowered to account for the number of comparisons being performed). Additionally, the value of Cohen's *d* (i.e., an appropriate effect size for the comparison between two means, which should be used to accompany the reporting of *U* of Mann–Whitney and *t*-test results, and which rules of thumb for interpreting effect size suggest that $d = 0.2$ be considered a small effect size, $d = 0.5$ represents a medium effect size, and $d = 0.8$ a large effect size) was calculated to understand the strength of the difference between the experimental and control groups (hypothesis 1) and pre-test and post-test (hypothesis 2).

3. Results

Tests of the two hypotheses were conducted using Bonferroni adjusted alpha levels of 0.010 per test (0.05/5) (hypothesis 1) and 0.004 per test (0.05/14) (hypothesis 2).

Pertaining to hypothesis 1, after obtaining a non-normal distribution from the Kolmogorov–Smirnov test on school performance, the *U* of the Mann–Whitney analyses did not yield any statistically significant differences between the experimental and control group of NtN tutees on school performance in the instrumental and STEM subjects at the end of the school year (Table 2).

Table 2. Intergroup (experimental vs. control) comparisons of NtN tutees' school performance.

Subject/Group	N	Mean	Standard Deviation	Mann–Whitney *U* Test	*p*-Value	*d* of Cohen
Natural Science						
Experimental	25	7.40	1.68	267.00	0.37	0.26
Control	25	6.96	1.74			
Social Science						
Experimental	25	7.20	1.73	277.00	0.48	0.22
Control	25	6.80	1.85			
Spanish Language and Literature						
Experimental	25	6.96	1.64	306.50	0.90	0.05
Control	25	7.04	1.72			
Mathematics						
Experimental	25	7.08	1.52	264.00	0.33	0.29
Control	25	6.64	1.52			
First Foreign Language (English)						
Experimental	25	7.20	1.84	251.00	0.22	0.31
Control	25	6.64	1.72			

In relation to hypothesis 2, the Kolmogorov–Smirnov test yielded a normal distribution of the NtN tutors' pre-test and post-test average scores on personality and key competences for employment. The *t*-student dependent test did not reveal any statistically significant pre-test/post-test differences between the NtN tutors' results regarding the BIP scales [39] (Table 3).

Table 3. NtN tutors' pre-test/post-test comparisons of personality and key competencies for employment.

Scale/Phase	N	Mean	Standard Deviation	*t*-Student Dependent Test	*p*-Value	*d* of Cohen
Achievement motivation						
Pre-test	26	56.54	6.63	−0.48	0.63	0.07
Post-test	26	56.04	6.89			
Power motivation						
Pre-test	26	41.92	3.80	−0.99	0.33	0.23
Post-test	26	40.96	4.50			
Leadership motivation						
Pre-test	26	42.92	7.29	0.59	0.56	0.08
Post-test	26	43.61	9.26			
Conscientiousness						
Pre-test	26	62.42	4.93	−0.69	0.50	0.14
Post-test	26	61.58	6.95			
Flexibility						
Pre-test	26	45.31	7.11	−1.14	0.26	0.22
Post-test	26	43.69	7.73			
Action orientation						
Pre-test	26	54.38	8.43	−1.39	0.18	0.27
Post-test	26	52.04	9.02			
Social sensitivity						
Pre-test	26	55.85	4.98	0.17	0.87	0.03
Post-test	26	56.00	5.95			
Sociability						
Pre-test	26	55.77	7.79	−2.24	0.03	0.28
Post-test	26	53.19	10.53			
Openness to contact						
Pre-test	26	56.27	10.42	0.17	0.87	0.02
Post-test	26	56.46	10.55			
Team orientation						
Pre-test	26	50.50	7.40	−2.81	0.01	0.54
Post-test	26	46.15	8.74			
Assertiveness						
Pre-test	26	49.30	6.38	−0.14	0.89	0.02
Post-test	26	49.15	6.21			
Emotional stability						
Pre-test	26	51.77	8.42	−1.47	0.15	0.23
Post-test	26	49.50	11.11			
Work under pressure						
Pre-test	26	48.92	9.01	−0.72	0.48	0.11
Post-test	26	47.85	10.64			
Self-confidence						
Pre-test	26	48.85	9.43	0.10	0.92	0.01
Post-test	26	48.96	9.01			

4. Discussion

This study aimed to demonstrate the capacity of the NtN-Granada, firstly, to improve the academic performance amongst primary education students in the instrumental and STEM subjects through the IBSE teaching method with a strong hands-on component and activities designed ad hoc for this project, and secondly, to increase the quality of competency-based co-curricular learning experiences available for university students. Of course, all was all done in collaboration with the institutional managers and teachers of the participating educational centers and under the direction and supervision of the university staff responsible for the program. Therefore, considering the results obtained, we can

ascertain that the program has some promise, although more conclusive evidence for the program's impact requires additional replications over time and across schools. In particular, the following conclusions can be drawn: (1) since there were no statistically significant differences between the experimental and control groups of NtN tutees in terms of their school performance once the program was finished, we have rejected hypothesis 1, and (2) because there were no statistically significant pre-test/post-test differences between the average scores achieved by the NtN tutors in personality and key competences for the employment, we have also rejected hypothesis 2.

Given the particular purposes of the study, the results should be interpreted not only considering the probability and statistical significance values [46], but also, and perhaps more importantly, on the basis of the effect size [47]. In fact, effect size is a more appropriate statistic since it measures the direction and magnitude of a treatment effect and, unlike significance tests, this index is independent of sample size [48]. In this sense, regarding hypothesis 1, an interpretation of the effect size values obtained reveals that a hypothetical member of the experimental group of NtN tutees could achieve a school performance higher than that of 58% members of the control group. In fact, although the size of the effect that was achieved must be interpreted as small [48] or small to medium [49], certain authors state that a change higher than 0.10 points can be considered as a significant improvement, even more so if this improvement is uniform in all the members of the group and is maintained over time [47]. Moreover, specialized literature indicates that effect sizes of around 0.20 have more practical or clinical significance when they are based on measures of outcomes that are more difficult to change (e.g., school performance) [50,51]. Therefore, the results of the program seem to confirm the initially proposed estimates about the effects of improvement that participation in this type of program entails for school performance of the primary education students in the STEM subjects [12,15,16,25–27,52–59]. Alternatively, the lack of statistically significant effects on school performance may be related to the problem of underpowered treatment (i.e., the results from short treatment schedules, poor attendance by participants, fragmented/poor implementation, and uninvolved parents) [28]. For example, tutoring sessions began two months after the course started (i.e., early September) due to the program preparation arrangements (e.g., dissemination, sampling). This delay led to a situation in which NtN tutees of the experimental group were close to completing the first three-month evaluation period, with very limited exposure to supporting activities of the program and their consequences on school grades for that evaluation period.

The need to terminate the intervention period before the NtN tutors started their own exam period at the university (i.e., early June) not only limited the final number of tutoring sessions but also prevented the possibility of delivering key tutoring sessions during June, which was also the final exam and evaluation period for NtN tutees of the experimental group. In this regard, and from a comparative perspective, the NtN-Rutgers activity not only lasted for almost a whole year but also delivered between three and four sessions per week. Thus, an intervention of sufficient duration ensuring maximum dosage would be necessary for future implementations, as several authors recommend [60,61].

Although many contents from the curriculum (e.g., biology, physics, earth sciences, technology, and mathematics) were addressed by the program activities, the mix of students from fifth and sixth graders in the experimental group made it difficult to fully align the session contents (e.g., natural sciences) with their morning classes in all cases, which might have conditioned the results.

Moreover, the behavior of some NtN tutees of the experimental group (e.g., lateness, attendance at sessions without necessary school materials, level of enthusiasm) and NtN tutors (e.g., delays recording the tutoring sessions in the workbooks and difficulties in carrying out the needs analysis, as well as in the design and execution of the tutoring sessions, sometimes due to the lack of prior preparation or level of enthusiasm) could underlie the limited effects of the program [26].

Insufficient planning and coordination structure between tutors and teachers to align the activities displayed in the regular classroom and those delivered in the afterschool support activities (e.g., delays in communication through email) could have also hampered the productivity of the program (effectiveness) [62]. Thus, the limited capacity of the program based on the IBSE teaching method to

deliver the promised learning objectives and competencies may have been mediated, on the one hand, by the lack of changes in the teaching methodologies affecting the majority of Spanish primary and secondary teachers, including science teachers [63], and, on the other hand, by the use of inadequate tests to assess and measure the kind of knowledge and competencies delivered by the IBSE teaching method, which should have formally been incorporated in the last 10 years in Spain, as has happened in many European countries, around key competences [64]. Particularly, one of the seven key competences in the Spanish case, that is, mathematical competence and basic competences in science and technology, including such components of scientific competence, such as evaluating and designing scientific inquiry or interpreting data and evidence scientifically derived from the PISA framework [65], could greatly benefit from the IBSE activities of the program. Nevertheless, such a gain in competences has not been measured by the type of exams implemented, which, consequently, is not mirrored in the final scores.

Additionally, insufficient follow-up and supervision by those responsible for the program (e.g., time cost associated with the program monitoring sessions in two schools for four days per week) due to lack of resources (e.g., lack of time) may also help explain the final results.

In relation to the program effects on NtN tutors, regardless of the effect size accomplished, participation in the program did not have the forecasted effects on NtN tutors' development of personality and key competencies for employment. The results obtained did not align with those found by the specialized literature on the benefits of student participation in experiences that include S-L and/or peer tutoring in their intervention design [29–31,33,36].

These discrepancies may be attributed to the response shift bias. This threat to internal validity takes place when the participant's metric for answering questions from pre-test to post-test changes due to a new understanding of the concept being investigated [30,66]. In this regard, the NtN tutors had a high perception of their own competence level in the pre-test, but the quasi-professional experiences they went through, the problems and difficulties they had to cope with, and the practice of professional behavior (e.g., participation in decision-making processes, data recording, monitoring and evaluation), increased their awareness and knowledge of the pitfalls of their profession, affecting their post-test responses to these evaluation materials. Inquiry experiences are not very common in Spanish primary science teachers training programs, and poor knowledge of inquiry amongst the programs' students is common [21], which makes the use of these co-curricular S-L activities attractive for high quality training purposes. Additionally, future research should focus on the expected professional outputs linked to IBSE interventions on the tutors' attitudes to science, particularly if relevance and novelty for the pre-service teachers, as is the case in this study, is involved [67].

Limitations

The findings of this study have to be interpreted with caution due to some limitations inherent to our non-probabilistic sampling and correlated samples design. For example, our non-probabilistic sampling entailed some limitations compromising the representativeness of the sample, and therefore the potential generalizations of research findings. However, this type of sampling is inevitable, especially when the resources, time, and/or workforce are limited [68].

In relation to the small sample size, the probability of finding an educational or clinically relevant difference that is statistically significant is low (type II error) and the chance of inconclusive results is high [69].

In this regard, when the power of the NtN-Granada was calculated, current sample-size flat rules of thumb for using a sample size of between 20–40 subjects per treatment arm in a two-armed trial [70] were followed. Unfortunately, although this study's matched pairs design intended to reduce error variability by equating both groups on key variables, the lack of randomisation to conform both groups did not guarantee that all relevant potential confounding variables are equated among our treatment levels prior to intervention.

5. Conclusions

This study's results have contributed to determining the feasibility of the components of the NtN-Granada (i.e., protocol, recruitment of subjects, measurement instruments, data collection and analysis), identifying their weaknesses, establishing the needed modifications to planning and designing a larger-scale study, and demonstrating its potential effectiveness [71]. In fact, preliminary data revealed a small to medium impact on the school performance of primary education students. Therefore, curriculum materials provided by NtN-Rutgers (e.g., lesson plans, lecture notes, activity work sheets, design of hands-on experiments, and instruments for data collection that were adapted to the Spanish educational system), as well as guidance on NtN-Granada design, implementation, and evaluation decisions, have definitely influenced how Spanish STEM curriculum was delivered. Additionally, this experience seems particularly relevant and useful to university students. Therefore, the NtN-Granada's program intends to ensure inclusive and equitable quality education and promote lifelong learning opportunities for primary education students and university students. Moreover, S-L experiences enable university social responsibility to be incorporated as a core of education for sustainability, and some of the scientific-technological activities of the program were oriented to environmental issues and sustainable education [72,73].

Furthermore, the promising results obtained with this experience could be radically improved with higher levels of involvement, interest, and collaboration between the different educational agents that participated in the different phases of the program, as other studies also pointed out [26], as well as with more systemic support [59]. The allocation of basic economic resources to strengthen management functions and performance before, during, and after the life of the program will benefit the fidelity of the program as well as the impact of the intervention on those key dependent variables identified in both samples (e.g., short-term outcome), but also on medium-term outcomes (e.g., attitudes towards inquiry and STEM, and attitudes towards more socially inclusive teaching). Those financial resources would also help extend the program beyond the school year into the summertime when children are free from school duties and responsibilities and have more time to learn from alternative educational resources.

References

1. European Centre for the Development of Vocational Training. *Statistics and Indicators. Rising STEMs*; European Commission: Brussels, Belgium, 2014.
2. National Science Board. *Science and Engineering Indicators 2016*; National Science Foundation: Arlington, VA, USA, 2016.
3. European Centre for the Development of Vocational Training. *Skills Panorama. Inspiring Choices on Skills and Jobs in Europe*; European Commission: Brussels, Belgium, 2014.
4. Caprile, M.; Palmén, R.; Sanz, P.; Dente, G. *Encouraging STEM Studies in the Labour Market*; European Parliament: Brussels, Belgium, 2015. [CrossRef]
5. Shapiro, H.; Fuglsang, S.; Frøhlich, K. *Does the EU Need More STEM Graduates? Final Reports*; European Commission: Brussels, Belgium, 2015. [CrossRef]

6.   Ministerio de Educación, Cultura y Deporte, Spanish Ministry of Education, Culture and Sport. *Anuario de Indicadores Universitarios*; University Indicators Yearbook; Secretaría General de Universidades: Madrid, Spain, 2017.

7.   Chang, M.J.; Sharkness, J.; Hurtado, S.; Newman, C.B. What Matters in College for Retaining Aspiring Scientists and Engineers from Underrepresented Racial Groups. *J. Res. Sci. Teach.* **2010**, *51*, 555–580. [CrossRef]

8.   Marginson, S.; Tytler, R.; Freeman, B.; Roberts, K. *STEM: Country Comparisons. International Comparisons of Science, Technology, Engineering and Mathematics (STEM) Education*; Final Report; Australian Council of Learned Academies: Melbourne, Australia, 2013.

9.   Wang, X. Why students choose STEM Majors: Motivation, High School Learning, and Postsecondary Context of Support. *Am. Educ. Res. J.* **2013**, *50*, 1081–1121. [CrossRef]

10.   United Nations. *Transforming our World: The 2030 Agenda for Sustainable Development*; United Nations Department of Economic and Social Affairs: New York, NY, USA, 2015.

11.   González, H.B.; Kuenzi, J.J. *Science, Technology, Engineering, and Mathematics (STEM) Education: A Primer*; Congressional Research Service: Washington, DC, USA, 2012.

12.   An, D. A Meta-Analysis of the Effectiveness of STEM-Programs in the United States. Ph.D. Thesis, Cleveland State University, Cleveland, OH, USA, 2013.

13.   Forsthuber, B.; Motiejunaite, A.; De Almedia, A.S. *Science Education in Europe: National Policies, Practices and Research*; Eurydice: Brussels, Belgium, 2011. [CrossRef]

14.   Kearney, C. *Efforts to Increase Students´ Interest in Pursuing Science, Technology, Engineering and Mathematics Studies and Careers*; National Measures Taken by 30 Countries—2015 Report; European Schoolnet: Brussels, Belgium, 2016.

15.   Van den Hurk, A.; Meelissen, M.; Van Langen, A. Intervention in Education to Prevent STEM Pipeline Leakage. *Int. J. Sci. Educ.* **2018**, *41*, 150–164. [CrossRef]

16.   Young, J.; Ortiz, N.; Young, J. STEMulating interest: A Meta-Analysis of the Effects of Out-Of-School Time on Students STEM Interest. *IJEMST* **2017**, *5*, 62–74. [CrossRef]

17.   Bybee, R.W. *The Case for STEM Education. Challenges and Opportunities*; National Science Teachers Association: Arlington, VI, USA, 2013.

18.   Breiner, J.M.; Johnson, C.C.; Harkness, S.S.; Koehler, C.M. What is STEM? A Discussion about Conceptions of STEM in Education and Partnerships. *Sch. Sci. Math.* **2012**, *112*, 3–11. [CrossRef]

19.   European Commission. *Science Education for Responsible Citizenship*; Publications Office of the European Union: Luxembourg, 2015. [CrossRef]

20.   Rocard, M.; Csermely, P.; Jorde, D.; Lenzen, D.; Walberg, H.; Hemmo, V. *Science Education Now: A Renewed Pedagogy for the Future of Europe*; European Commission: Brussels, Belgium, 2007.

21.   García, A.; Criado, A.M.; Cruz, M. Primary Pre-Service Teachers' Skills in Planning a Guided Scientific Inquiry. *Res. Sci. Educ.* **2017**, *47*, 989–1010. [CrossRef]

22.   Minner, D.D.; Levy, A.J.; Century, J. Inquiry-Based Science Instruction-What is it and does it Matter? Results from a Research Synthesis Years 1984 to 2002. *J. Res. Sci. Teach.* **2010**, *47*, 474–496. [CrossRef]

23.   Lazonder, A.W.; Harmsen, R. Meta-Analysis of Inquiry-Based Learning: Effects of Guidance. *Rev. Educ. Res.* **2016**, *86*, 681–718. [CrossRef]

24.   Olson, S.; Loucks, S. *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*; National Academies Press: Washington, DC, USA, 2000.

25.   Camasso, M.J.; Jagannathan, R. Improving Academic Outcomes in poor Urban Schools Through Nature-Based Learning. *Camb. J. Educ.* **2017**, *48*, 263–277. [CrossRef]

26.   Jagannathan, R.; Camasso, M.J.; DelaCalle, M. The Effectiveness of a Head-Heart-Hands Model for Natural and Environmental Science Learning in Urban Schools. *Eval. Program. Plann.* **2018**, *66*, 53–62. [CrossRef]

27.   Camasso, M.J.; Jagannathan, R. Nurture thru Nature: Creating Natural Science Identities in Populations of Disadvantaged Children through Community Education Partnership. *J. Environ. Educ.* **2017**, *49*, 30–42. [CrossRef]

28.   Hayward, M.S. Service Learning in the Community College: Drive to Employment. *Community Coll. J. Res. Pract.* **2014**, *38*, 838–841. [CrossRef]

29.   Westover, J.H. *Academic Service-Learning across Disciplines. Models, Outcomes and Assessment*; Common Ground Publishing: Champaign, IL, USA, 2012.

30. Arco, J.L.; Fernández, F.D.; Fernández, J.M. The Impact of a Peer-Tutoring Program on Quality Standards in Higher Education. *High. Educ.* **2011**, *62*, 773–788. [CrossRef]

31. Arco, J.L.; Fernández, F.D.; Hervás, M. Evidence-Based Peer-Tutoring Program to Improve Students' Performance at the University. *Stud. High. Educ.* **2019**. [CrossRef]

32. Clayton, P.H.; Bringle, R.G.; Hatcher, J.A. *Research on Service Learning. Conceptual Frameworks and Assessment*; Stylus Publishing: Sterling, VI, USA, 2013.

33. Fernández, F.D.; Arco, J.L.; Hervás, M.; Delgado, L.C. Transformar la Educación Superior y Obligatoria a través de Aprendizaje-Servicio y Mentoría [Transform Higher and Compulsory Education via Service-Learning and Mentoring]. *Univ. Psychol.* **2018**, *17*, 1–12. [CrossRef]

34. Tindall, J.A.; Black, D.R. *Peer Programs: An In-Depth Look at Peer Helping*, 2nd ed.; Routledge: New York, NY, USA, 2009.

35. Topping, K.J. Peer Tutoring: Old Method, New Developments. *Infanc. Aprendiz.* **2015**, *38*. [CrossRef]

36. Yorio, P.L.; Ye, F. A Meta-Analysis of the Effects of Service-Learning on the Social, Personal, and Cognitive Outcomes of Learning. *Acad. Manag. Learn. Educ.* **2012**, *11*, 9–27. [CrossRef]

37. Bringle, R.G.; Hatcher, J.A.; Jones, S.G. *International Service Learning. Conceptual Frameworks and Research*; Stylus Publishing: Sterling, VI, USA, 2011.

38. Patton, M.Q. *Qualitative Research and Evaluation Methods*, 3rd ed.; Sage Publications: Thousand Oaks, CA, USA, 2002.

39. Hossiep, R.; Paschen, M. *Business-Focused Inventory of Personality*; Hogrefe: Göttingen, Germany, 1998.

40. Ato, M.; López, J.J.; Benavente, A. Un Sistema de Clasificación de los Diseños de Investigación en Psicología [A Classification System of Research Designs in Psychology]. *An. Psicol.* **2013**, *29*, 1038–1059. [CrossRef]

41. Fernández, M.; Mena, L.; Riviere, J. *Fracaso y Abandono Escolar en España*; School Failure and Drop-Out in Spain; Fundación La Caixa: Barcelona, Spain, 2010.

42. Rivkin, S.G.; Hanushek, E.A.; Kain, J.F. Teachers, Schools, and Academic Achievement. *Econometrica* **2005**, *73*, 417–458. [CrossRef]

43. Mildenhall, P.; Cowie, B.; Sherriff, B. A STEM Extended Learning Project to Raise Awareness of Social Justice in a Year 3 Primary Classroom. *Int. J. Sci. Educ.* **2018**, *41*, 471–489. [CrossRef]

44. National Research Council. *Successful K-12 STEM Education. Identifying Effective Approaches in STEM*; The National Academies Press: Washington, DC, USA, 2011.

45. Yager, R.E.; Brunkhorst, H. *Exemplary STEM Program: Design for Success*; National Science Teachers Association: Arlington, VA, USA, 2014.

46. Goldschmid, M.L. Teaching and Learning in Higher Education: Recent Trends. *High. Educ.* **1976**, *5*, 437–456. [CrossRef]

47. Coe, R. It´s the Effect Size, Stupid. What Effect Size is and Why it is Important. In Proceedings of the British Educational Research Association Annual Conference, University of Exeter, Exeter, UK, 12–14 September 2002.

48. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1988.

49. Sawilowsky, S.S. New Effect Size Rules of Thumb. *J. Mod. Appl. Stat. Methods* **2009**, *8*, 597–599. Available online: https://digitalcommons.wayne.edu/cgi/viewcontent.cgi?article=1536&context=jmasm (accessed on 5 August 2020). [CrossRef]

50. Durlak, J.A. How to Select, Calculate and Interpret Effect Sizes. *J. Pediatr. Psychol.* **2009**, *34*, 917–928. [CrossRef] [PubMed]

51. Hedges, L.V.; Hedberg, E.C. Intraclass Correlation Values for Planning Group-Randomized Trials in Education. *Educ. Eval. Policy Anal.* **2007**, *29*, 60–87. [CrossRef]

52. Afterschool Alliance. *STEM Learning in Afterschool: An Analysis of Impacts and Outcomes*; Afterschool Alliance: Washington, DC, USA, 2011.

53. Becker, K.; Park, K. Effects of Integrative Approaches Among Science, Technology, Engineering, and Mathematics (STEM) Subjects on Students´ Learning: A Preliminary Meta-Analysis. *J. STEM Educ.* **2011**, *12*, 23–36. Available online: https://www.jstem.org/jstem/index.php/JSTEM/article/download/1509/1394 (accessed on 5 August 2020).

54. Dabney, K.P.; Tai, R.H.; Almarode, J.T.; Miller-Friedmann, J.L.; Sonnert, G.; Sadler, P.M.; Hazari, Z. Out-Of-School Time Science Activities and their Association with Career Interest in STEM. *Int. J. Sci. Educ.* **2012**, *2*, 63–79. [CrossRef]

55. Denson, C.D.; Hailey, C.; Stallworth, C.A.; Householder, D.L. Benefits of Informal Learning Environments: A Focused Examination of STEM-Based Program Environments. *J. STEM Educ.* **2015**, *16*, 11–15. Available online: https://www.jstem.org/jstem/index.php/JSTEM/article/download/1893/1630 (accessed on 5 August 2020).

56. Hudson, P.; English, L.; Dawes, L.; King, D.; Baker, S. Exploring Links Between Pedagogical Knowledge Practices and Student Outcomes in STEM Education for Primary Schools. *Aust. J. Teach. Educ.* **2015**, *40*. [CrossRef]

57. Kong, X.; Dabney, K.P.; Tai, R.H. The Association between Science Summer Camps and Career Interest in Science and Engineering. *Int. J. Sci. Educ.* **2014**, *4*, 54–65. [CrossRef]

58. Sahin, A.; Ayar, M.C.; Adiguzel, T. STEM Related After-School Program Activities and Associated Outcomes on Student Learning. *Educ. Sci. Theory Pract.* **2014**, *14*, 309–322. [CrossRef]

59. Shah, A.M.; Wylie, C.; Gitomer, D.; Noam, G. Improving STEM Program Quality in Out -of School- Time: Tool Development and Validation. *Sci. Educ.* **2018**, *102*, 38–259. [CrossRef]

60. Levine, P.B.; Zimmerman, D.S. *Targeting Investments in Children: Fighting Poverty When Resources are Limited*; University of Chicago Press: Chicago, IL, USA, 2010.

61. Tough, P. *How Children Succeed: Grit, Curiosity, and the Hidden Power of Character*; Houghton Mifflin Harcourt: Boston, MA, USA, 2013.

62. Grineski, S.; Daniels, H.; Collins, T.; Morales, D.X.; Frederick, A.; Garcia, M. The Conundrum of Social Class: Disparities in Publishing among STEM Students in Undergraduate Research Programs at a Hispanic Majority Institution. *Sci. Educ.* **2018**, *102*, 283–303. [CrossRef] [PubMed]

63. Mateos, A.; García, B.; Bejarano, M.T. How Spanish Science Teachers Perceive the Introduction of Competence-Based Science Teaching. *J. Balt. Sci. Educ.* **2016**, *15*, 371–381. Available online: http://www.scientiasocialis.lt/jbse/?q=node/509 (accessed on 5 August 2020).

64. Pepper, D. Assessing Key Competences across the Curriculum—And Europe. *Eur. J. Educ.* **2011**, *46*, 335–353. [CrossRef]

65. Organization for Economic Co-operation and Development. PISA 2015 Science framework. In *PISA 2015 Assessment and Analytical Framework*; OECD Publishing: Paris, France, 2017. [CrossRef]

66. Howard, G.S. Response-Shift Bias. A Problem in Evaluating Interventions with Pre/Post Self-Reports. *Eval. Rev.* **1980**, *4*, 93–106. [CrossRef]

67. Palmer, D.H.; Dixon, J.; Archer, J. Identifying Underlying Causes of Situational Interest in a Science Course for Pre-Service Elementary Teachers. *Sci. Educ.* **2016**, *100*, 1039–1061. [CrossRef]

68. Etikan, I.; Musa, S.A.; Alkassim, R.S. Comparison of Convenience Sampling and Purposive Sampling. *AJTAS* **2016**, *5*. [CrossRef]

69. Machin, D.; Campbell, M.J.; Tan, S.B. *Sample Size Tables for Clinical Studies*, 3rd ed.; Wiley Blackwell: Chichester, UK, 2008.

70. Whitehead, A.L.; Julious, S.A.; Cooper, C.L.; Campbell, M.J. Estimating the Sample Size for a Pilot Randomised Trial to Minimise the Overall Trial Sample Size for the External Pilot and Main Trial for a Continuous Outcome Variable. *Stat. Methods Med. Res.* **2016**, *25*, 1057–1073. [CrossRef]

71. Bobronnikov, E.; Sahni, S.D.; Bozzi, L. *A Guide for Reporting on Evaluations for the US Department of Education Mathematics and Science Partnerships (MSP)*; Abt Associates Inc.: Cambridge, MA, USA, 2013.

72. Cabedo, L.; Royo, M.; Moliner, L.; Guraya, T. University Social Responsibility towards Engineering Undergraduates: The Effect of Methodology on a Service-Learning Experience. *Sustainability* **2018**, *10*, 1823. [CrossRef]

73. Toronjo-Hornillo, L.; Del Castillo-Andrés, Ó.; Campos-Mesa, M.D.C.; Díaz Bernier, V.M.; Zagalaz Sánchez, M.L. Effect of the Safe Fall Programme on Children's Health and Safety: Dealing Proactively with Backward Falls in Physical Education Classes. *Sustainability* **2018**, *10*, 1168. [CrossRef]