The impact of teaching style and FCI gain on the performance in mechanics tests

C Azzopardi¹, L Camilleri² and C Bonello³

¹ Physics, University of Malta, Junior College
² Statistics and Operations Research, University of Malta
³ Mathematics and Science Education, University of Malta
carmel.g.azzopardi@um.edu.mt

Abstract. The purpose of this study is to identify the relationship, if any, between the gain in the force concept inventory (FCI) and study level, gender, repeaters/non-repeaters, teaching style and the end-of-year mechanics result. The study focussed on all students (N = 395) admitted to a post-secondary institution in Malta in October 2015 and who participated in the pretest and posttest FCI exercise, studying physics at advanced (N = 162) or intermediate level (N = 233). There were 36 students at advanced level and 44 students at intermediate level who did not participate in the posttest FCI. In this study a multiple-choice test was used comprising of 30 questions about mechanics concepts and a questionnaire (by Centre for Occupational Research and Development (CORD)) was given to the lecturers to identify the teaching goals and methods used. The pretest was given on the second week of the academic year, October 2015, so that their response was not influenced by mechanics teaching and the posttest (identical to pretest) was given in the end of the scholastic year, May 2016. The results of the study suggest that only 13.7% of the participants exceeded the average-expected gain of 24%. The gain is higher for males, for participants studying advanced level physics and for participants who were taught by particular lecturers with particular teaching goals and methods. There is a positive correlation between the gain in the FCI and the end-of-year score in mechanics questions about forces and Newton’s laws of motion.

1. Introduction

After a three-year course in physics between the age of 14 and 16, successful candidates may opt to further their studies in the subject at advanced or intermediate level. Azzopardi [1] explains that students study Newton’s laws of motion, linear motion, energy, projectiles in vertical motion and conservation of linear momentum in collisions at second level, but not circular motion and vector addition at an angle. When students are admitted to this post-secondary institution, they opt to study physics at advanced or intermediate level. The intermediate level is designed to have a weight of one-third the commitment and effort of an advanced level subject. At the end of the first year of the course, candidates take a school-based exam to determine their fate for the second year of their course.

Misconceptions or preconceptions are stumbling blocks in the learning process, as described by Halloun and Hestenes [2]. They categorised misconceptions about motion into the principles and influences about it. This study paved the way to the design of the force concept inventory (FCI) to check the level of misconceptions in candidates. Maizam and Norsuzana [3] argued that matching teaching
and learning styles is beneficial in successful learning outcomes; however, it is more relevant to developmental learners than highly-motivated mature learners.

2. Objective
The purpose of this quantitative study was to identify whether misconceptions in mechanics were reduced by the learning experience in the first year of the course. The study was designed to identify the relationships, if any, that exist between the gain on the FCI and gender, study level, repeating and non-repeating students, and teaching styles. It was also used to identify if the gain in FCI has an effect on the performance in mechanics questions asked in the end-of-first-year exam.

3. Design
The Force Concept Inventory (FCI) was originally published in 1992 [4], and revised in 1995 [5]. The FCI consists of multiple-choice questions with one Newtonian answer and four plausible non-Newtonian options. It measures six dimensions of force concept in kinematics, Newton’s three laws of motion, net force, and kinds of force.

A questionnaire set by CORD (Centre for Occupational Research and Development) was given to the teachers responsible for teaching the mechanics course to all classes in the department. The questionnaire has 12 sections, each carrying four statements describing teaching goals and methods. In each section the teacher was requested to rank the four statements from 1 (most describing) to 4 (least describing).

4. Research Questions
The research was meant to address the following questions:

a) Is the gain in FCI related to gender, course level, repeating/non-repeating students and teaching styles?
b) Does the posttest score indicate effective teaching methods?
c) Are the gain in FCI and the performance in the mechanics questions in the end-of-year exam correlated?

5. Methods
This research is based on a multiple-choice-question test given to the students, a questionnaire about teaching goals and methods given to the teachers, and the end-of-year-school-based exam.

The FCI (30 multiple-choice questions) was given as a pretest at the beginning of the first year \( (N = 475) \) and the same test was given as a posttest at the end of the same scholastic year \( (N = 408) \). The percentage gain for each student who has participated in both tests \( (N = 395) \) was calculated as follows:

\[
\text{Gain} = \left( \frac{\%\text{posttest} - \%\text{pretest}}{100 - \%\text{pretest}} \right) \times 100\%
\]

Hestenes and Halloun Error! Reference source not found. argued that when the FCI is administered as a pretest and posttest, the gain can be individually computed for each student. This gain reveals ‘high gain’ if it is greater than 60% and ‘low gain’ if it is less than 20%. Under conventional instruction it is expected to be around 25% of the maximum possible gain.

A questionnaire about teaching goals and methods was completed by all teaching staff who were delivering the mechanics content throughout the year. The teacher ranked from 1 (most) to 4 (least) describing his/her goals and methods according to written statements. The replies were computed and plotted on two charts for teaching goals and methods respectively. The teaching goals are described into four categories namely (1) rote-abstract; (2) rote-applied; (3) understanding-abstract and (4) understanding-applied. The teaching methods are also described into four other categories namely (1) symbolic-individual; (2) symbolic-cooperative groups; (3) enactive-individual and (4) enactive-cooperative groups. After computation, each teacher was assigned a code 1, 2, 3, or 4 for his/her teaching
methods and goals. However, some scores were assigned the codes 12, 13, 24 and 34 to illustrate in-between categories; see figures 1 and 2.

In the end of the first-year instruction, the department sets a 3-hour paper for both advanced and intermediate level groups. The questions are about mechanics, waves and optics. In this research, only the sum of marks scored for mechanics questions were collected and analysed.

SPSS version 21 was used for statistics analysis. The Kolmogorov-Smirnov and Shapiro-Wilk test were used to assess the normality assumption of the percentage score distribution. To compare mean percentage scores between several independent groups, the One-Way ANOVA test was used when the score distribution was Normal; whereas its non-parametric alternatives (Kruskal-Wallis test) was used when the score distribution was not Normal. To compare mean pretest and posttest performance scores, the Paired-Sample t-test was used when the score distribution was Normal; while its non-parametric alternative (Wilcoxon-signed-ranks test) was used when the score distribution was not Normal. Both p-values of the parametric (P) and non-parametric (N-P) tests are displayed when the score distribution of one variable is normal and the other is not normal. Descriptive Statistics tables display mean, maximum and minimum scores. Finally the Pearson (parametric) and Spearman-rho (non-parametric) correlation tests were used to determine whether the relationship between the mechanics and FCI-gain scores was significant.

6. Results
All results hereunder were derived for the following groups: all participants together, advanced and intermediate level groups. There were 395 participants who participated in both pretest and posttest. There were 8 (2.03%) participants in the pretest and no participants in the posttest who left out more than six questions. However, there were 291 (73.67%) participants in the pretest and 318 (80.51%) participants in the posttest who answered all the questions completely. The mean percentage score in the pretest was 26.60% (s.d. = 10.43) and in the posttest it was 32.33% (s.d. = 13.42). According to Hestenes [6], in their research in autumn 2000, among 200 high school teachers and 20000 students beginning high school physics, showed an average percentage score of 26% in the pretest and 42% in the posttest following traditional teaching. In this research, the pretest average percentage score is close to 26% but the posttest average percentage score was much lower than 42%.

This study finds that 1 (0.25 %) is a ‘stronger’ student (gain > 60%) and 323 (81.77%) are ‘weaker’ students (gain < 20%) according to Hestenes [5]. 104 (26.33%) of the ‘weaker’ students obtained a negative gain, which means that they performed worse in the posttest than in the pretest.

6.1. Is the gain in FCI related to gender, repeating/non-repeating students and teaching styles?
There were 217 (54.9%) male and 178 (45.1%) female participants all together. In the intermediate level group, there were 98 (42.1%) male participants and 135 (57.9%) female participants. In the advanced level group there were 119 (73.5%) male participants and 43 (26.5%) female participants. Table 1 shows that male participants obtained a significantly higher mean percentage score in the pretest, posttest and the gain in both levels of instruction. Only the percentage gain at the intermediate level was not significantly higher for male participants.

There were 365 (92.4%) non-repeating (NR) and 30 (7.6%) repeating (R) participants in the entire cohort. In the intermediate level group, there were 223 (95.7%) non-repeating and 10 (4.3%) repeating participants. In the advanced level group there were 142 (87.7%) non-repeating and 20 (12.3%) repeating participants. Table 2 shows that there is a significant difference in the mean score percent in the pretest and posttest for the advanced level group but no difference for the intermediate group. Although the table shows a slightly higher mean score percent in favour of the repeating group, the maximum score percent in pretest, posttest and gain is always higher for the non-repeating group. This analysis shows that although the mean score percent is slightly higher for the repeating group, the maximum score percent is obtained by the non-repeating group.
Table 1. Mean percentage scores for male and female participants (N-P – Kruskal-Wallis; P – ANOVA).

| Group       | Max Score % (Male) | Mean Score % (Male) | Max Score % (Female) | Mean Score % (Female) | Difference in mean score in % | Significant | N-P p-value | P p-value |
|-------------|--------------------|---------------------|----------------------|-----------------------|-------------------------------|-------------|-------------|-----------|
| All         |                    |                     |                      |                       |                               |             |             |           |
| Pretest     | 66.67              | 29.68               | 50.00                | 22.85                 | 6.83                          | YES         | 0.000       | -         |
| Posttest    | 80.00              | 36.88               | 66.67                | 26.78                 | 10.1                          | YES         | 0.000       | -         |
| Gain        | 62.50              | 9.89                | 52.17                | 4.50                  | 5.39                          | YES         | 0.001       | -         |
| Intermediate|                    |                     |                      |                       |                               |             |             |           |
| Pretest     | 50.00              | 27.69               | 50.00                | 22.67                 | 5.02                          | YES         | 0.000       | -         |
| Posttest    | 70.00              | 31.26               | 56.67                | 22.78                 | 8.48                          | YES         | 0.001       | -         |
| Gain        | 43.75              | 4.39                | 43.48                | 3.30                  | 1.09                          | NO          | -           | 0.570     |
| Advanced    |                    |                     |                      |                       |                               |             |             |           |
| Pretest     | 66.67              | 31.32               | 46.67                | 23.41                 | 7.91                          | YES         | 0.000       | -         |
| Posttest    | 80.00              | 41.51               | 66.67                | 29.92                 | 11.59                         | YES         | 0.000       | -         |
| Gain        | 62.50              | 14.42               | 52.17                | 8.29                  | 6.13                          | YES         | 0.020       | 0.036     |

Table 2. Mean percentage scores for repeating (R) and non-repeating (NR) participants (N-P – Kruskal-Wallis; P – ANOVA).

| Group       | Max Score % (Repeating) | Mean Score % (Repeating) | Max Score % (non-repeating) | Mean Score % (non-repeating) | Difference in % | Significant | N-P p-value | P p-value |
|-------------|-------------------------|--------------------------|----------------------------|----------------------------|----------------|-------------|-------------|-----------|
| All         |                         |                          |                            |                            |                |             |             |           |
| Pretest     | 46.67                   | 32.67                    | 66.67                      | 26.10                     | 6.57           | YES         | 0.002       | -         |
| Posttest    | 63.33                   | 39.89                    | 80.00                      | 31.71                     | 8.18           | YES         | 0.004       | -         |
| Gain        | 45.45                   | 10.93                    | 62.50                      | 7.18                      | 3.75           | NO          | 0.318       | -         |
| Intermediate|                         |                          |                            |                            |                |             |             |           |
| Pretest     | 46.67                   | 28.33                    | 50.00                      | 24.62                     | 3.71           | NO          | 0.466       | -         |
| Posttest    | 50.00                   | 30.67                    | 70.00                      | 27.97                     | 2.70           | NO          | 0.390       | -         |
| Gain        | 21.74                   | 2.91                     | 43.75                      | 3.80                      | -0.89          | NO          | -           | 0.850     |
| Advanced    |                         |                          |                            |                            |                |             |             |           |
| Pretest     | 46.67                   | 34.83                    | 66.67                      | 28.43                     | 6.40           | YES         | 0.008       | -         |
| Posttest    | 63.33                   | 44.50                    | 80.00                      | 37.58                     | 6.92           | YES         | 0.026       | -         |
| Gain        | 45.45                   | 14.94                    | 62.50                      | 12.94                     | 2.00           | NO          | 0.484       | 0.536     |

There were 9 classes at intermediate and 8 classes at advanced level physics. These classes were coached by 11 teachers, two of which were teaching the advanced level classes, seven of which were teaching the intermediate level classes and two were teaching both levels. Table 3 indicates that the best performance was obtained by the advanced level group C10 (posttest and gain) while the intermediate group C5 obtained the worst average gain.

Table 4 illustrates that teacher T1 obtained a significantly higher maximum mean score in the posttest and gain among all participants in both levels. Teacher T1 was coaching groups C14 and C15, both advanced level groups. Ironically these groups (C14 and C15) did not obtain the highest mean score in Table 3. Among intermediate level groups, teacher T9 obtained a significantly higher maximum score in the posttest. T9 was coaching the intermediate group C2 and is indicated in Table 3 to obtain a significantly high mean score in the posttest.
Table 3. Mean percentage scores for participants in different classes (N-P – Kruskal-Wallis; P – ANOVA).

| Group   | Maximum mean score in % | Minimum mean score in % | Difference in % | Significant | N-P p-value (0.05) | P p-value (0.05) |
|---------|-------------------------|-------------------------|------------------|-------------|--------------------|-----------------|
| All     |                         |                         |                  |             |                    |                 |
| Pretest | 32.50                   | C12                     | 20.58            | C7          | 11.92              | NO              | 0.059           | -               |
| Posttest| 43.57                   | C10                     | 22.75            | C7          | 20.82              | YES             | 0.000           | -               |
| Gain    | 23.35                   | C10                     | -2.97            | C5          | 20.38              | YES             | 0.000           | -               |
| Intermediate | 26.67                | C9                      | 20.58            | C7          | 6.09               | NO              | 0.216           | -               |
| Posttest| 31.67                   | C2                      | 22.75            | C7          | 8.92               | YES             | 0.004           | -               |
| Gain    | 9.29                    | C2                      | -2.97            | C5          | 12.26              | NO              | -               | 0.075           |
| Advanced|                         |                         |                  |             |                    |                 |                 |
| Pretest | 32.50                   | C12                     | 26.67            | C17         | 5.83               | NO              | 0.000           | -               |
| Posttest| 43.57                   | C10                     | 33.50            | C11         | 10.07              | NO              | 0.221           | -               |
| Gain    | 23.35                   | C10                     | 4.42             | C13         | 18.93              | YES             | 0.001           | 0.001           |

Table 4. Mean percentage scores for participants grouped by their tutor (T1 – T11).

| Group   | Maximum mean score in % | Minimum mean score in % | Difference in % | Significant | N-P p-value (0.05) | P p-value (0.05) |
|---------|-------------------------|-------------------------|------------------|-------------|--------------------|-----------------|
| All     |                         |                         |                  |             |                    |                 |
| Pretest | 29.38                   | T5                      | 20.58            | T10         | 8.80               | YES             | 0.032           | -               |
| Posttest| 42.07                   | T1                      | 22.75            | T10         | 19.32              | YES             | 0.000           | -               |
| Gain    | 19.65                   | T1                      | -2.97            | T3          | 22.62              | YES             | 0.000           | -               |
| Intermediate | 26.67                | T11                     | 20.58            | T10         | 6.09               | NO              | 0.216           | -               |
| Posttest| 31.67                   | T9                      | 22.75            | T10         | 8.92               | YES             | 0.004           | -               |
| Gain    | 9.29                    | T9                      | -2.11            | T3          | 11.40              | NO              | -               | 0.100           |
| Advanced|                         |                         |                  |             |                    |                 |                 |
| Pretest | 31.50                   | T5                      | 27.86            | T2          | 3.64               | NO              | 0.623           | -               |
| Posttest| 42.07                   | T1                      | 34.70            | T2          | 7.37               | NO              | 0.083           | -               |
| Gain    | 19.65                   | T1                      | 6.89             | T5          | 12.76              | YES             | 0.000           | 0.002           |

The teaching goals were organised in two dimensions on a grid ranging from rote (R) to understanding (U) on the y-axis and abstract (AB) to applied (AP) on the x-axis as shown in Figure 1. The response of each teacher was assigned values whose computation results in a number 1, 2, 3 or 4. In one case, the response gave 24 which meaning that there is no distinction between rote and understanding (R/U) goals.

Figure 1. Score description for teaching goals.

Table 5 indicates that all teachers who taught advanced level groups were all of the same understanding-applied goals. For the intermediate level groups, the understanding-abstract goal gave a significant high maximum score in the gain compared to rote/understanding-applied goals but not in the pretest and posttest scores. The understanding-applied goal strategy gave a significant high score in posttest and gain when the cohort was considered.
Table 5. Mean percentage scores for participants grouped by their teacher’s goals (N-P – Kruskal-Wallis; P – ANOVA).

| Group               | Maximum score in % | Minimum score in % | Difference in % | Significant | N-P p-value (0.05) | P p-value (0.05) |
|---------------------|--------------------|--------------------|-----------------|-------------|--------------------|-----------------|
| All Pretest         | 27.47              | U-AP               | 24.21           | R-AP        | 3.26               | NO              | 0.325           | -               |
| All Posttest        | 34.51              | U-AP               | 24.84           | R/U-AP      | 9.18               | YES             | 0.000           | -               |
| All Gain            | 9.59               | U-AP               | -2.97           | R/U-AP      | 12.56              | YES             | 0.002           | -               |
| Intermediate Pretest| 25.89              | U-AB               | 24.21           | R-AP        | 1.68               | NO              | 0.635           | -               |
| Intermediate Posttest| 31.45             | U-AB               | 24.84           | R/U-AP      | 6.61               | NO              | 0.128           | -               |
| Intermediate Gain   | 6.80               | U-AB               | -2.97           | R/U-AP      | 9.77               | YES             | -               | 0.036           |
| Advanced Pretest    | SAME TEACHING GOALS, UNDERSTANDING-APPLIED FOR ALL TEACHERS |
| Advanced Posttest   | SAME TEACHING GOALS, UNDERSTANDING-APPLIED FOR ALL TEACHERS |
| Advanced Gain       | SAME TEACHING GOALS, UNDERSTANDING-APPLIED FOR ALL TEACHERS |

The teaching methods were organised in two dimensions on a grid ranging from symbolic (S) to enactive (E) on the y-axis and individual (I) to group (G) on the x-axis as shown in Figure 2. The response of each teacher was assigned values whose computation results in a number 1, 2, 3 or 4. There are two cases where the response gave 34 which means that there is no distinction between individual and group (I/G) while the other response gave 13 which means that there is no distinction between symbolic and enactive (S/E) methods.

Figure 2. Score description for teaching methods.

Table 6 shows that the symbolic/enactive-individual teaching method gave a significant high score in the posttest and gain among the intermediate group but not for the advanced level group.

6.2. Does the posttest score indicate effective teaching methods?

The Wilcoxon signed-ranks test or/and the paired-sample t-test were used to determine if the teachers and their teaching methods were effective in the posttest scores depending on the distributions of the pretest and posttest scores.

Table 7 illustrates that teachers T3, T6, T7, T10, T11 and the symbolic/enactive-group teaching method were not effective in the teaching instruction.
### Table 6. Mean percentage scores for participants grouped by their teacher’s methods (N-P – Kruskal-Wallis; P – ANOVA).

| Group   | Maximum score in % | Minimum score in % | Difference in % | Significant | N-P p-value (0.05) | P p-value (0.05) |
|---------|---------------------|---------------------|-----------------|-------------|--------------------|-----------------|
| All     |                     |                     |                 |             |                    |                 |
| Pretest | 28.41 S/E-I         | 25.06 S-I           | 3.35            | NO          | 0.750              | -               |
| Posttest| 36.51 S/E-I         | 24.84 E-I/G         | 11.67           | YES         | 0.000              | -               |
| Gain    | 11.22 S/E-I         | -2.97 E-I/G         | 14.19           | YES         | 0.000              | -               |
| Intermediate |               |                     |                 |             |                    |                 |
| Pretest | 26.40 S/E-I         | 23.26 E-I           | 3.14            | NO          | 0.363              | -               |
| Posttest| 31.55 S-I           | 24.84 E-I/G         | 6.71            | YES         | 0.007              | -               |
| Gain    | 8.01 S-I            | -2.97 E-I/G         | 10.98           | YES         | -                  | 0.008           |
| Advanced |                 |                     |                 |             |                    |                 |
| Pretest | 29.74 S/E-I         | 27.86 E-G           | 1.88            | NO          | 0.747              | -               |
| Posttest| 39.91 S/E-I         | 34.70 E-G           | 5.21            | NO          | 0.192              | -               |
| Gain    | 13.65 E-I           | 9.50 E-G            | 4.15            | NO          | 0.209              | 0.351           |

### Table 7. Effectiveness of teachers and teaching methods on the posttest scores (N-P – Wilcoxon; P – Paired sampled t-test).

| Group   | Number | Mean score % | St. Dev | Min score % | Max score % | N-P p-value (0.05) | P p-value (0.05) | Significant |
|---------|--------|--------------|---------|-------------|-------------|--------------------|-----------------|-------------|
| T1      | 45     | 27.93        | 8.14    | 13.33       | 43.33       | 0.000              | 0.000           | EFFECTIVE   |
| Pretest |        | 42.07        | 14.43   | 13.33       | 70.00       |                    |                 |             |
| Posttest|        | 34.70        | 14.34   | 6.67        | 46.67       | 0.001              | 0.001           | EFFECTIVE   |
| T2      | 39     | 25.91        | 11.25   | 3.33        | 50.00       | -                  | 0.840           | NOT EFFECTIVE |
| Pretest |        | 24.84        | 9.38    | 6.67        | 46.67       |                    |                 |             |
| Posttest|        | 31.55        | 14.22   | 13.33       | 76.67       | 0.000              | -               | EFFECTIVE   |
| T3      | 63     | 28.41        | 11.65   | 10.00       | 66.67       | 0.000              | -               | EFFECTIVE   |
| Pretest |        | 36.51        | 14.22   | 13.33       | 76.67       |                    |                 |             |
| Posttest|        | 29.38        | 12.51   | 10.00       | 63.33       | 0.000              | -               | EFFECTIVE   |
| T4      | 64     | 25.00        | 10.02   | 3.33        | 50.00       | -                  | 0.326           | NOT EFFECTIVE |
| Pretest |        | 27.36        | 10.86   | 3.33        | 50.00       |                    |                 |             |
| Posttest|        | 22.00        | 7.45    | 0.00        | 33.33       | 0.223              | 0.247           | NOT EFFECTIVE |
| T5      | 24     | 25.89        | 9.42    | 10.00       | 43.33       | 0.011              | 0.023           | EFFECTIVE   |
| Pretest |        | 31.45        | 10.89   | 13.33       | 56.67       |                    |                 |             |
| Posttest|        | 30.39        | 9.96    | 6.67        | 46.67       | 0.001              | 0.000           | EFFECTIVE   |
| T6      | 28     | 24.17        | 9.96    | 6.67        | 46.67       | 0.001              | 0.000           | EFFECTIVE   |
| Pretest |        | 31.67        | 10.32   | 13.33       | 56.67       |                    |                 |             |
| Posttest|        | 20.58        | 8.74    | 6.67        | 50.00       | 0.246              | -               | NOT EFFECTIVE |
| T7      | 23     | 22.75        | 11.22   | 10.00       | 63.33       | 0.696              | 0.479           | NOT EFFECTIVE |
| Pretest |        | 26.67        | 8.17    | 13.33       | 46.67       |                    |                 |             |
| Posttest|        | 28.41        | 9.37    | 16.67       | 46.67       | 0.000              | 0.000           | EFFECTIVE   |
| T8      | 58     | 25.06        | 9.64    | 6.67        | 46.67       | 0.000              | 0.000           | EFFECTIVE   |
| Pretest |        | 31.55        | 10.53   | 13.33       | 56.67       |                    |                 |             |
| Posttest|        | 27.35        | 10.98   | 6.67        | 63.33       | 0.000              | 0.000           | EFFECTIVE   |
| T9      | 132    | 25.68        | 9.05    | 0.00        | 50.00       | -                  | 0.001           | EFFECTIVE   |
| Pretest |        | 34.80        | 14.64   | 10.00       | 70.00       |                    |                 |             |
| Posttest|        | 29.52        | 12.38   | 3.33        | 80.00       | 0.000              | -               | EFFECTIVE   |
| T10     | 63     | 28.41        | 11.65   | 10.00       | 66.67       | 0.000              | -               | EFFECTIVE   |
| Pretest |        | 36.51        | 14.22   | 13.33       | 76.67       |                    |                 |             |
| Posttest|        | 35.91        | 11.25   | 3.33        | 50.00       | 0.000              | 0.580           | NOT EFFECTIVE |
| T11     | 31     | 24.84        | 9.38    | 6.67        | 46.67       | -                  | 0.580           | NOT EFFECTIVE |
6.3. Are the gain in FCI and the performance in the mechanics questions in the end-of-year exam correlated?

At the end of their first year, the participants take a school-based exam to qualify for the continuation of their second year. This exam has questions about mechanics, optics and waves. There were 388 participants who completed the pretest, posttest and the end-of-year exam. 232 were intermediate level and 156 were advanced level students. For this correlation analysis, only those questions specifically about forces and Newton’s laws of motion were considered. The percentage mark for the mechanics score was computed on the mechanics questions attempted by the candidate. The Spearman-rho and the Pearson correlation test were implemented for the entire, advanced and intermediate groups depending on whether the dependent mechanics scores were normally distributed or not.

The maximum percentage score in mechanics questions in the end-of-year exam was 96.7 in the advanced level group and 65.0 in the intermediate group.

Table 8. Correlation coefficient between the FCI gain and Mechanics score (N-P – Spearman; P – Pearson).

| Group   | Number | Spearman Coefficient | Pearson Coefficient | Significant | N-P p-value (0.05) | P p-value (0.05) |
|---------|--------|-----------------------|----------------------|-------------|--------------------|-----------------|
| All     | 388    | 0.315                 | -                    | YES         | 0.000              | -               |
| Intermediate | 232 | 0.181                 | -                    | YES         | 0.006              | -               |
| Advanced | 156    | -                     | 0.381                | YES         | -                  | 0.000           |

Figure 3. FCI gain against mechanics score for all students (N = 388).

Figure 4. FCI gain against mechanics score for intermediate level students (N = 232).

Figure 5. FCI gain against mechanics score for advanced level students (N = 156).

7. Discussion and conclusions

The research shows that all participants obtained an average score of 26.73% in the pretest, 32.29% in the posttest and an overall gain of 4.87%. The average pretest score was within the expected margin of 26% while the posttest score was much lower than the expected value 42% for traditional teaching [6]. Students with more than 60% of the maximum possible gain were termed ‘stronger students’ while students with less than 20% of the maximum possible gain were termed ‘weaker students’. In this research, 9 individuals were ‘strong’ and 284 were ‘weak’, leaving 102 individuals to lie in between.

For all 395 participants, the study found that male participants performed better than female participants on all counts except for the gain in the intermediate level. Repeating participants did not perform significantly better than non-repeating ones except in the pretest, posttest for all participants and the advanced level group. The maximum percentage score in pretest, posttest and gain were obtained by the non-repeating group. Class C10 (advanced) performed better in posttest and gain scores while C2 (intermediate) in the gain. Teachers T1 (teaching 45 students) obtained the highest mean in the posttest among advanced level group while T9 (teaching 28 students) got the highest mean in the gain score among the intermediate group. In general the teaching goals did not make a difference to the scores except for understanding-applied goal in the posttest and gain by all participants, and the understanding-
abstract goal in the gain by intermediate participants. The symbolic-individual teaching method made a significant difference on the posttest and gain results obtained by all intermediate level participants. However, the study shows that the teaching methods did not show a significant difference among the advanced level participants. Teachers T1, T2, T4, T5, T8 and T9 were found to be effective in the posttest results while the teaching methods symbolic-individual, enactive-individual, enactive-group and symbolic/enactive-individual were effective for the posttest results.

There is a positive correlation between the gain obtained in the FCI and the mechanics score obtained by the participants in the end-of-year results. When zooming in to advanced/intermediate level, teachers/teaching styles, the results can be different than the general picture.

Acknowledgment
I would like to thank Prof. Liberato Camilleri for his advice on statistical tests performed in this study, Dr. Charles Bonello for his valuable guidance, Dr. Russell Mizzi (subject coordinator) and Mr Paul Xuereb (the principal) for the permission to carry this research at the institution, the physics staff who participated in the administration of the survey and the participants.

References
[1] Azzopardi C 2016 The Misconceptions in Mechanics among students after completing their Secondary Level Education in Malta. In: Pixel (ed) International Conference, The Future of Education 2016, Conference Proceedings (Florence, Italy 30 June - 1 July 2016) p. 511. Florence: Libreria universitaria.
[2] Halloun I A and Hestenes D 1985 Common sense concepts about motion Am. J. Phys. 53(11), 1056–65. DOI: 10.1119/1.14031
[3] Maizam A and Norsuzana Z 2008 Methods of teaching and goals of teaching: Teaching styles of teachers in higher institutions. Retrieved January 3, 2016, from http://eprints.uthm.edu.my/2277/
[4] Hestenes D, Wells M and Swackhamer G 1992 Force Concept Inventory Phys. Teach. 30 141–58. DOI:10.1119/1.2343497
[5] Hestenes D and Halloun I 1995 Interpreting the Force Concept Inventory: A response to March 1995 critique by Huffman and Heller Phys. Teach. 33 502–06. DOI:10.1119/1.2344278
[6] Hestenes D 2016 ASU Modeling Instruction & MNS Degree Programs: legacy site. Retrieved March 1, 2016, from Research & evaluation: http://modeling.asu.edu/R&E/Research.html