The Effect of Engineering Education, on Spatial Ability, in Virtual Environments

Tibor Guzsvinecz[1], Éva Orbán-Mihálykó[2], Erika Perge[3], Cecília Sik-Lányi[4]

[1]Department of Information Technology and its Applications, Faculty of Information Technology, University of Pannonia, Gasparich Márk utca 18/A, 8900 Zalaegerszeg, Hungary

[2]Department of Mathematics, Faculty of Information Technology, University of Pannonia, Egyetem utca 10, 8200 Veszprém, Hungary

[3]Department of Basic Technical Studies, Faculty of Engineering, University of Debrecen, Ótemető utca 2, 4028 Debrecen, Hungary

[4]Department of Electrical Engineering and Information Systems, Faculty of Information Technology, University of Pannonia, Egyetem utca 10, 8200 Veszprém, Hungary

[1]guzsvinecz.tibor@zek.uni-pannon.hu, [2]orbane@almos.uni-pannon.hu, [3]perge@eng.unideb.hu, [4]lanyi@almos.uni-pannon.hu

Abstract: Since spatial skills are important in the modern world, tests that can improve them are included in the curriculum of engineering studies. In this article, the results of a pre-test and a post-test are presented, which were performed in the herein developed, virtual environment. The former was conducted in the beginning, while the latter was done at end of semester 2019/2020/1, with 240 engineering students. The spatial skills of students were measured on these tests using the Mental Rotation Test (MRT), Mental Cutting Test (MCT), and the Purdue Spatial Visualization Test (PSVT). Between the pre-test and post-test, spatial ability enhancing courses were attended by the students. The results show that their performance in the post-test, notably improved – albeit slightly. This improvement differs between various user groups and test types: regarding the students, 1st-year, right-handed or those who were at most 18 years old improved significantly, between the two tests; regarding the test types, significant improvement only occurred during the two MRT tests. While no such difference was detected on the pre-test, civil engineering students significantly outperformed mechanical engineering students in the post-test, while males who scored at least 50 points on the post-test, performed significantly better than females on the pre-test.

Keywords: cognitive skills; education; mental rotation test; mental cutting test; purdue spatial visualization test; spatial ability; virtual environments; virtual reality

1 Introduction

As several jobs concerning engineering require well-developed spatial skills, they are essential in the modern age. For example, architects, interior designers, engineers and technicians should excel in spatial aptitude tasks. Due to the importance of spatial skills, it was suggested by Miller that spatial ability training should be included in the curriculum of engineering studies [1]. In the study of Miller and Bertoline, it was also concluded that these skills can be improved through time as they are not biological susceptibilities [2]. Through the years, a considerable number of paper-based tests was developed to enhance the spatial skills of the users, such as the three that are focused on in this paper: the Mental Rotation Test (MRT) [3], the Mental Cutting Test (MCT) [4] and the Purdue Spatial Visualization Test (PSVT) [5].

Since the world is transitioning into a digital one, new versions of these spatial ability improving tests are recreated in virtual or in augmented environments [6-9], although their numbers are scarce. Some studies even include gamification elements in the tests with the goal to make the tests feel like a game by introducing levels, points, and rewards to motivate the user [10] [11]. Also, it became easier to create digital versions of these tests and to measure the spatial skills of the users since the inception of the field of Cognitive InfoCommunications (CogInfoCom) [12-15]. As the CogInfoCom environment consists of multiple fields (such as education [16], human-computer interaction [17-20], connection between eye-tracking and complex cognitive processes [21] [22], sense of presence [23]) and the user is placed into virtual reality (VR) – or into virtual laboratories [24] – the role of visuality in virtual space [25-27], and smart environments [28], new cognitive capabilities can emerge [29]. Thus, with CogInfoCom, it is possible to investigate how human cognitive capabilities can be merged and extended with the cognitive capabilities of digital devices to provide easier interaction between humans and other, artificially cognitive agents.

Besides these, research of spatial skills is also part of this environment [30-32]. According to Kovari et al., a connection exists between problem solving [33], Education 4.0 and CogInfoCom [34] [35]. This means that Education 4.0, by its concept, encourages the implementation and development of individual knowledge and skills using modern digital educational technology systems and individualized education.

To investigate this connection further, the authors of this paper would like to investigate whether traditional engineering education can affect the spatial skills of the users in virtual environments. For example, the use of VR can even improve real-world spatial skills [36]. However, what about tasks in reality? Can these real-world tasks improve the spatial skills of people in virtual environments? Thus, it would be interesting to see whether education of students during their courses can affect their spatial skills in VR. This is crucial to be investigated, since a VR system is complex, made up of various parts [37] and the users are just as essential as the other parts [38-40].

2 Research Questions and Hypotheses

To reach this goal, two tests were taken by 240 engineering students: one in the beginning of the 2019/2020/1 semester and another one in the end of it. These two tests will be referred to as pre-test and post-test, respectively. On both tests, the spatial skills of students were measured using the MRT, MCT, and PSVT tests. Between these two tests, the students took spatial ability training courses. These courses lasted 14 weeks. During the practical sessions, there were groups of 30 students, and the sessions lasted 3 hours/week/group.

During these courses, students learnt the foundations of axonometry and perspective, the Monge two-plane representation, issues of spatial shapes' visibility, pruning tasks, the transformation of image plane systems, representing circles, forms of rotations, perpendicularity, matching tasks, and metric problems. Also, the development of spatial skills was helped by geometric bodies made from cardboard and 3D printing, as well as the various body models and animations created using a 3D modeling software. Taking these spatial forms in hand, rotating them, observing their properties, cutting the forms made of foam (Styrofoam), building from different forms, and taking them apart contributed to the acquisition of personal experience.

As can be seen, these courses in engineering education aim to enhance the spatial skills of students, therefore it could be interesting to see whether there are improvements between the results of the pre-test and post-test.

Therefore, after the pre-test concluded, the authors set up two research questions (RQ) about the post-test which are the following:

- RQ1: Did the results of the students significantly improve between the two tests after completing the spatial ability training courses?
- RQ2: Do the improvements vary among the different user groups?

The authors and teachers expected that improvement can be realized. To check this hope and get the answers, the same number of null hypotheses (Hs) were formed.

- H1: The results of the students do not significantly improve between the two tests.
- H2: The improvements do not vary among the different user groups.

3 Methodology

A spatial ability measuring virtual environment was developed in 2019 by the authors. This virtual environment can be used with a desktop display and the Gear VR head-mounted display. The former version uses Windows, while the latter uses Android. For this study however, only the former version was used. In this virtual environment, the spatial skills of the users could be measured with three types of tests: the MRT, MCT, and PSVT. When entering the virtual environment, the users are placed into a space and the selected test type appears in front of them. The users cannot change location in this space. They have to select their answers using a keyboard and/or mouse on PC. When using the Android version, they have to use the touchpad on the right side of the Gear VR to select their answers. Examples of these three test types in the virtual environment are shown in Figure 1.

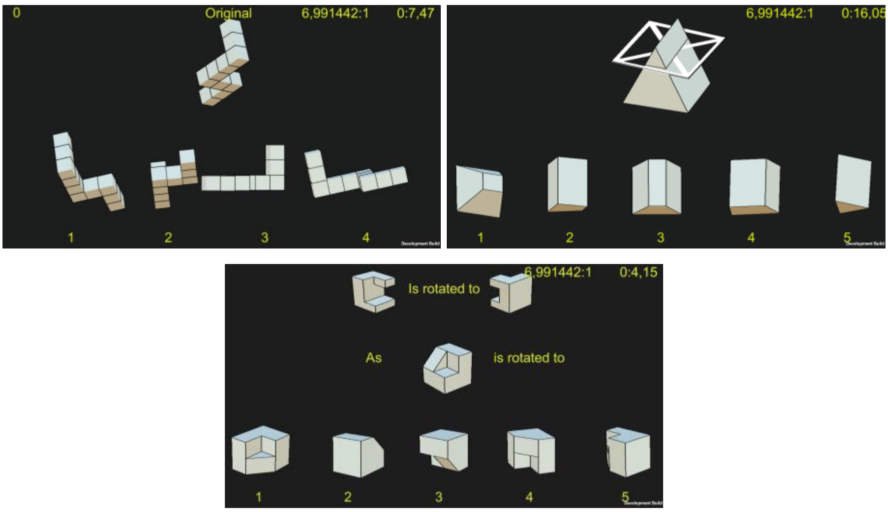

Figure 1

Examples of the three test types: MRT (left), MCT (right), and PSVT (center)

Each test type has ten questions and has to be done by the users three times. This means that 9 test sequences exist and a total of 90 answers is gathered from one student. However, the test questions were randomized in each case to prevent the students from remembering them. The completion times are also logged after each sequence. This measurement method was the same on both the pre-test and the post-test. There were no differences. Information regarding the users is logged as well by the application: when starting the virtual environment, the gender, primary hand, age, studies of the user and the years spent at the University, have to be entered into the application.

During the analysis of the post-test's results, a problem arose regarding the mentality of the students. Both the pre-test and the post-test were included in their university courses; therefore, it was mandatory for them to take the tests. In the case of the pre-test, the students assumed that their results could affect their grades on exams. However, the post-test was done after the students completed their exams, meaning that they realized that their answers do not influence their grades. Due to this, a possibility exists that some students answered the questions on the spatial ability tests very quickly. This changed the data distributions between the two tests. This can be seen in Figure 2.

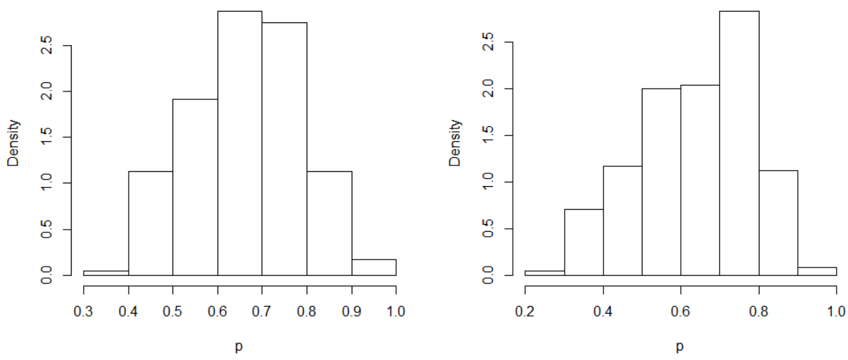

Figure 2

Distributions of the rates of correct answers in case of the pre-test (left) and post-test (right)

The Kolmogorov-Smirnov test was performed on each data frame. The results were the following: p-value = 0.6335 in the case of the pre-test and it is 0.03801 in the case of the post-test. This means that the hypothesis of normal distribution is accepted regarding the former, while it is rejected regarding the latter. However, the dispersion increased between the tests (from 0.1185 to 0.1423), while the average decreased between the two tests (from 0.6660 to 0.6391). A two-sample Welch-test was also done to test the equality of the expected values. According to the results (p-value = 0.02503), the two data frames significantly differ from each other. Even the rates of correct answers are worse in the case of the post-test, which is very-very surprising. It may suggest that some students did not take the test seriously. To investigate this suspicion further, the test completion times were investigated by the authors. This can be seen in Figure 3.

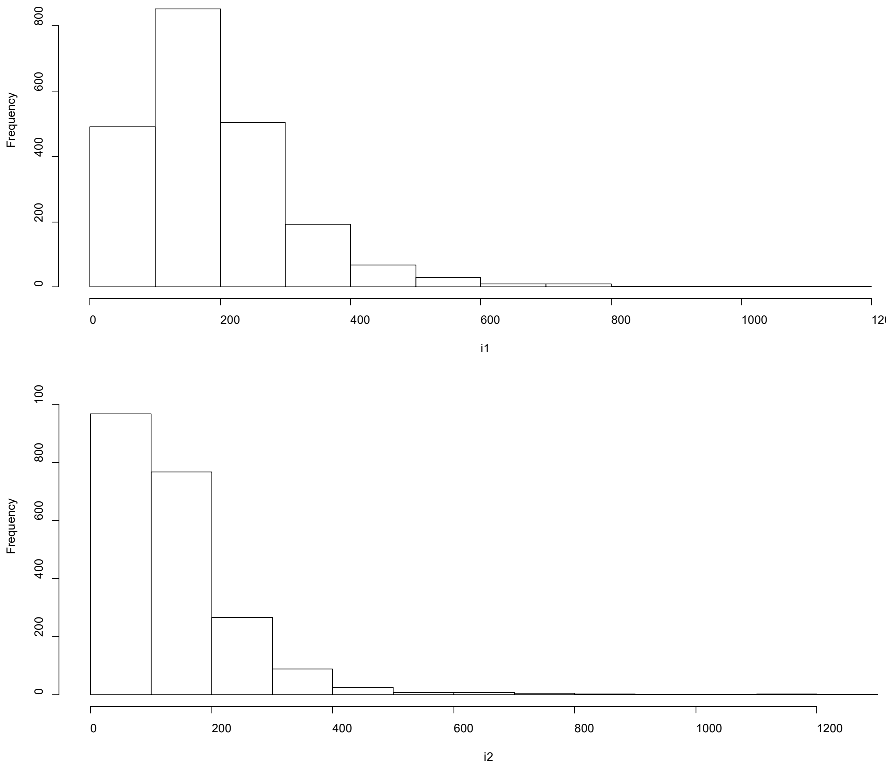

Figure 3
Histograms of the completion times regarding the pre-test (upper) and post-test (lower)

At the first sight, one can realize big differences between the two histograms: in the case of the post-test, there are extremely many small values among the completion times. To avoid this inaccuracy of "first sight", the authors analyzed the completion times data in detail. First, the descriptive statistics are the following: in the case of the pre-test, the smallest completion time is 7.9 seconds, the largest is 1168.4 seconds, the average is 189.3 seconds and the dispersion is 121.2 seconds, while they are 6.6 seconds, 1239.10 seconds, 134.61 seconds, and 115.5 seconds, respectively, in the case of the post-test. As can be seen, the average completion times are quite smaller in the case of the post-test.

A two-sample Welch-test was also done to see whether the completion times of the two tests are significantly different from each other. According to the results of the test (p-value $< 2.2\times10^{-16}$), they are. In Figure 4 their cumulative distribution functions are presented.

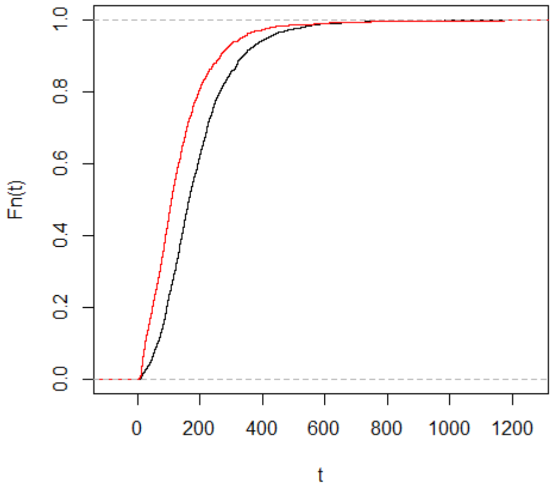

Figure 4
The cumulative distribution functions regarding the completion times of the pre-test (black) and post-test (red)

As can be seen in Figure 4, the red line (which belongs to the post-test) starts to increase earlier than the black one (which is the pre-test). This fact can also be observed from the kernel function. This means that in the case of the post-test, the number of small completion times has been increased which is possibly due to the "mandatory spatial ability tests" not influencing the grades of the students.

When talking about the completion times, the previous fact is also supported by the quantiles in the data: the authors calculated the deciles of data of the completion times and the results are presented in Table 1.

Table 1
The deciles of data of the completion times

Test	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Pre	7.9	64.4	94.1	119.5	142.9	165.1	193.7	223.2	264.0	337.6
Post	6.6	25.1	48.2	71.5	89.7	109.4	131.2	158.6	195.0	256.4

As can be seen in Table 1, approximately 20% of the data (431 records) regarding completion times are under 48.2 seconds in the case of the post-test. This is different in the case of the pre-test, as only 6% of data (131 records) are under 48.2 seconds. Naturally, there is a small possibility that this decrease in time is due to the students became too good at solving these tests. However, this is a very large decrease, moreover the probability of correct answers would have increased. This fact and the large increase in small times suggest otherwise.

Naturally, the data had to be carefully filtered to eliminate the answers of the students who did not take the tests seriously. For this, the test types were taken by the authors who tried to correctly solve the spatial ability questions as quickly as possible. The results show that their completion times were longer than a minute in each case. They even reached the two-minute mark in a few cases. Therefore, based on the authors' time and the data in the previous subsection, it is unrealistic to correctly solve a test type in under a minute. However, to give the benefit of the doubt, the authors concluded to filter the data with a 50-second limit. Of course, both the pre-tests and post-test data were handled in the same manner.

First, the authors had to decide whether the data or the students themselves are trustworthy. As the authors think that this property is the consequence of an attitude, a decision was made to eliminate people with all their data. Before filtering, the authors split the students into two groups: "trustworthy" and "untrustworthy" ones. Trustworthy students were the ones whose completion times were equal or longer than 50 seconds in at least 8 test sequences out of 9 (as each test type was done three times). Naturally, the definition of untrustworthy students is the complement of the trustworthy ones. Then all data of untrustworthy students were excluded. After making these two groups, the distribution of post-test's data changed which can be seen in the lower half of Figure 5.

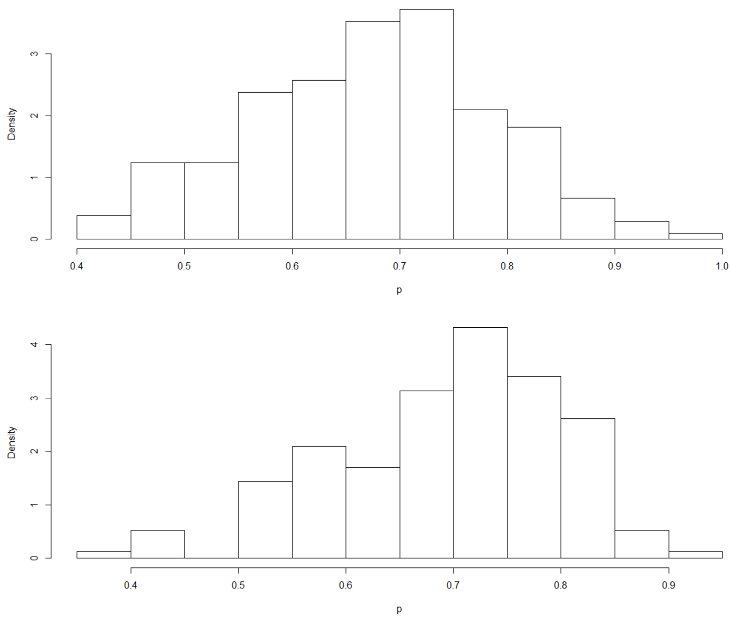

Figure 5
The filtered data distributions of the pre-test (upper) and the post-test (lower)

For sake of correct data analysis, the data of the pre-test and post-test were also filtered. According to the p-value = 0.8426 and p-value = 0.2072, the hypotheses of normal distributions were accepted in the case of the pre-test and post-test, respectively.

4 Results and Discussion

There are three subsections in this section. The general results are presented in the first. The hypothesis testing can be found in the second one. Lastly, the limitations of the study are detailed in the last one.

4.1 General Results

Both tests were done by 240 students. Regarding the rates of correct answers, the minimum is 0.3083, the maximum is 0.9667, the average is 0.6660 and the dispersion is 0.1185 in the case of the pre-test, while they are 0.2583, 0.9167, 0.6391, and 0.1423, respectively in the case of the post-test. After the data were filtered, these numbers changed. Regarding the pre-test, 210 students' data were in the filtered data set: their minimum rate of correct answers is 0.408, their maximum is 0.9667, the average is 0.6791 and their dispersion is 0.1110. Regarding the post-test, 153 students were in the filtered data set: their minimum rate of correct answers is 0.3583, their maximum is 0.9167, their average is 0.6986 and their dispersion is 0.1043.

After the data were filtered, it could be seen that the numerical value of the average rates of correct answers is larger in the case of the post-test. According to the results of the two-sample Welch-test, p-value = 0.0437 which means that the results are significantly better – albeit only slightly – in the case of the post-test. In the following subsection, this fact is investigated in detail with the filtered data. Afterward, the analysis of the results began.

First, the results of the male and female students were assessed, for both tests. These results are shown in Table 2.

Table 2
The results of males and females on the two tests

	Number of students	Min	Max	Average	Dispersion
Pre-test, male	183	0.4083	0.9667	0.6919	0.1078
Post-test, male	142	0.3583	0.9167	0.7028	0.1043
Pre-test, female	27	0.4417	0.7833	0.5923	0.0963
Post-test, female	11	0.5250	0.7750	0.6447	0.0921

As can be observed in Table 2, the average rates of correct answers are numerically better in the case of the post-tests regarding both males and females. First, the rates of correct answers on the pre-test were investigated in the filtered data set. Based on the results of the two-sample Welch-test, males who achieved at least 50 correct answers (out of 120 which is the number of possible correct answers) performed significantly better than females on the pre-test (p-value = 1.773×10^{-5}).

The next part of the investigation consisted of examining the rates of correct answers on the post-test between the genders. Using a two-tailed Welch-test, males performed almost slightly significantly better than females on the post-test (p-value = 0.06888); although the number of female students is small. However, since the differences in the averages are quite large numerically, the authors wanted to check the results with a one-tailed Welch-test which resulted in p-value = 0.03444. Due to the probability value, a significant difference exists: males performed better on the post-test.

Afterward, the differences between the males' rates of correct answers on the two tests were investigated. Both the two-tailed and one-tailed Welch-tests did not result in significant differences (p-value = 0.3568 and p-value = 0.1784, respectively). Thus, there are improvements in the case of males, but they are not significant.

Lastly, a similar analysis was done in the case of females. Just by observing their numerical data, we can note that their improvement is larger than that of males. Due to the results of the one-tailed Welch-test, the improvement of females is almost significant on the 0.05 level (p-value = 0.06623). If this level were 0.1, then their improvements would be considered significant. The larger level of significance (the probability of type I. error) can be reasoned by the small number of data concerning females.

Next, the results of the right-handed and left-handed students were assessed on both tests. These results are presented in Table 3.

Table 3

The results of right-handed and left-handed students on the two tests

	Number of students	**Min**	**Max**	**Average**	**Dispersion**
Pre-test, right-handed	185	0.408	0.967	0.683	0.109
Post-test, right-handed	135	0.450	0.917	0.705	0.098
Pre-test, left-handed	25	0.442	0.883	0.651	0.123
Post-test, left-handed	18	0.358	0.833	0.654	0.140

The average rates of correct answers improved on the post-test in the case of both groups, but it is greater in the case of right-handed students. Afterward, four Welch-tests were performed, and their results are presented in Table 4.

According to Table 4 significant improvement exists between the results of the two tests in the case of right-handed students (p-value = 0.0317). The differences between the results of all other groups are not significant.

Table 4
The differences in the results of right-handed and left-handed students on the two tests

	Test stat.	Significance	Number of tails
Pre-test, right-handed & Pre-test, left-handed	1.2287	0.2290	2
Post-test, right-handed & Post-test, left-handed	1.4800	0.1550	2
Pre-test, right-handed & Post-test, right-handed	-1.8636	0.0317	1
Pre-test, left-handed & Post-test, left-handed	-0.0769	0.4696	1

The studies of the students on both tests were investigated next. The numerical results are presented in Table 5, while the results of the comparison are shown in Table 6. In both tables, architectural or civil engineering students are shortened to CE and mechanical engineering students to ME. Mechatronics students also joined the post-tests, although they were not present on the pre-test. They are abbreviated to MC in the two tables.

Table 5
The results of various engineering students on the two tests

	Number of students	Min	Max	Average	Dispersion
Pre-test, CE	58	0.4083	0.8500	0.6592	0.1039
Post-test, CE	26	0.4500	0.8500	0.6670	0.1127
Pre-test, ME	152	0.4417	0.9667	0.6867	0.1134
Post-test, ME	98	0.3583	0.9167	0.6990	0.1039
Post-test, MC	29	0.5250	0.8917	0.7259	0.0925

Table 6
The differences in the results of different engineering students on the two tests

	Test stat.	Significance	Number of tails
Pre-test, CE & Pre-test, ME	-1.670	0.098	2
Post-test, CE & Post-test, ME	-1.308	0.199	2
Post-test, CE & Post-test, MC	-2.103	0.041	2
Post-test, ME & Post-test, MC	-1.335	0.188	2
Pre-test, CE & Post-test, CE	-0.300	0.766	2
Pre-test, ME & Post-test, ME	-0.881	0.379	2

According to the results presented in Table 6, only one significant difference exists due to p-value = 0.041: there is a significant difference between the results of civil engineering students and mechatronics students on the post-test.

The next to investigate was the years spent at the university. The results of the students are presented in Table 7. It should be noted that someone mistakenly marked "not a student" on the post-test as every tester was a university student.

Table 7
The results of engineering students on the two tests grouped by university years

	Number of students	Min	Max	Average	Dispersion
Pre-test, 1st year	207	0.4083	0.9667	0.6798	0.1108
Pre-test, 2nd year	2	0.4500	0.6750	0.5625	0.1590
Pre-test, 3rd year	1	0.7583	0.7583	0.7583	NA
Post-test, 1st year	150	0.3583	0.9167	0.6993	0.1040
Post-test, 2nd year	1	0.5750	0.5750	0.5750	NA
Post-test, 3rd year	1	0.6000	0.6000	0.6000	NA
Post-test, not a student	1	0.8167	0.8167	0.8167	NA

As can be seen in Table 7, most of the students were in their 1st year. It is shown by the data in Table 7, that the average rate of correct answers of 1st-year improved between the two tests. However, is the difference significant? To answer this question, a comparison was done between the results. Due to the result of the one-tailed Welch-test, the difference is significant (p-value = 0.045). This means that the results on the post-test were improved significantly compared to the pre-test.

Next, the results were investigated by age groups, but first, the number of students by age has to be assessed. This is presented in Figure 6.

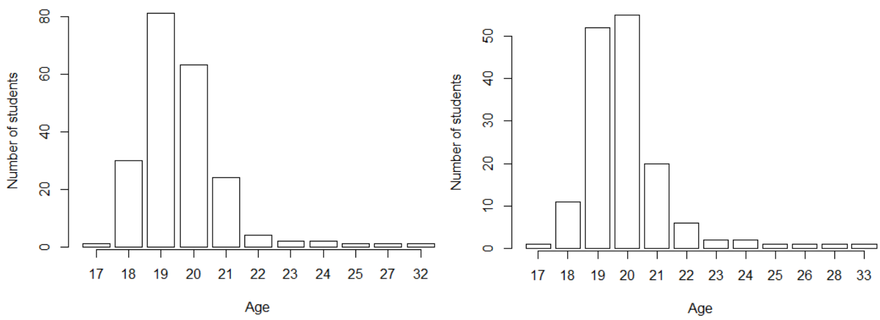

Figure 6
The number of students by age on the pre-test (left) and the post-test (right)

As can be seen in Figure 6, the two years which have the largest number of students are 19 and 20 in the case of both tests. In an earlier paper of the authors ([31]), two age groups were created: those students who are 18 years old at most and those who are over 18. Therefore, these two groups are used in this paper as well. The results of the users and the comparison can be seen in Tables 8 and 9, respectively.

Table 8
The results of the students by age groups on the two tests

	Number of students	Min	Max	Average	Dispersion
Pre-test, ≤ 18	31	0.4750	0.8583	0.6411	0.1099
Post-test, ≤ 18	12	0.4500	0.8167	0.6201	0.1060
Pre-test, > 18	179	0.4083	0.9667	0.6857	0.1105
Post-test, > 18	141	0.3583	0.9167	0.7053	0.1017

Table 9
The differences in the results of engineering students by age groups on the two tests

	Test stat.	Significance	Number of tails
Pre-test, ≤ 18 & Pre-test, > 18	-2.080	0.044	2
Post-test, ≤ 18 & Post-test, > 18	-2.681	0.019	2
Pre-test, ≤ 18 & Post-test, ≤ 18	0.576	0.570	2
Pre-test, > 18 & Post-test, > 18	-1.652	0.050	1

According to Table 9, a detectable difference exists between the two age groups' results on both tests (p-value = 0.044 and p-value = 0.019, respectively). No significant difference can be detected between the results of those who are ≤ 18 years old on both tests (p-value = 0.570), while a detectable difference exists between the results of those who are older than 18 on both tests (p-value = 0.050).

Lastly, the test types were assessed between the pre-test and post-test. The results grouped by the test types are presented in Table 10, while those of the comparisons are shown in Table 11.

Table 10
The results of the students by test types on the pre-test and post-test

	Number of students	Min	Max	Average	Dispersion
Pre-test, MRT	210	0.433	1.000	0.812	0.129
Pre-test, MCT	210	0.133	0.967	0.451	0.153
Pre-test, PSVT	210	0.133	0.967	0.640	0.172
Post-test, MRT	153	0.383	1.000	0.846	0.114
Post-test, MCT	153	0.133	0.833	0.454	0.150
Post-test, PSVT	153	0.033	1.000	0.648	0.185

According to the results of the Welch-test, each test type can be distinguished from each other on every level of significance. After comparing the results on the pre-test and post-test, it can be concluded that the only significant improvement can be found on the MRT test (p-value = 0.004). It can also be observed that in the case of the other two test types (where the average is drastically lower than on the MRT test), no significant improvement appears.

Table 11
The differences in the results of the students by test types on the pre-test and post-test

	Test stat.	Significance	Number of tails
Pre-test, MRT & Pre-test, MCT	26.153	0	2
Pre-test, MRT & Pre-test, PSVT	11.610	0	2
Pre-test, MCT & Pre-test, PSVT	-11.878	0	2
Post-test, MRT & Post-test, MCT	25.729	0	2
Post-test, MRT & Post-test, PSVT	11.283	0	2
Post-test, MCT & Post-test, PSVT	-10.065	0	2
Pre-test, MRT & Post-test, MRT	-2.629	0.004	1
Pre-test, MCT & Post-test, MCT	-0.189	0.425	1
Pre-test, PSVT & Post-test, PSVT	-0.407	0.342	1

4.2 Hypothesis Testing

Based on the results presented in the previous subsection, both null hypotheses (H1 and H2) are rejected. This means that education has an effect on the results and the alternative hypotheses became true in every case.

The first to discuss is H1 which originally states that "the results of the students do not significantly improve between the two tests". This is rejected as the results inside the two data frames (pre-test and post-test) were compared to each other and it yielded p-value = 0.0437. According to the previous probability value, the results are significantly different between the two tests, but this significance is only slight. This means that attending the spatial ability training subjects which focus on multiple types of spatial problems can increase these skills of the students. Therefore, what the students learn in a real environment can affect virtual ones as well.

Keep in mind that the two tests were conducted in the same semester. Therefore, there is a possibility that longer or future courses that focus on spatial skills can enhance them even more.

The next – and last – null hypothesis to discuss is H2. It originally stated that "the improvements do not vary between the different user groups". This was also rejected. For this, first, the results grouped by the students' gender were assessed. When comparing the pre-tests in the filtered data, males who achieved at least 50 correct answers performed significantly better than females on the pre-test (p-value = 1.773×10^{-5}). According to Table 2, males performed significantly better than females on the post-test by approximately 9.02% (p-value = 0.03444). It should be noted that this difference was approximately 16.82% between males and females on the pre-test, therefore engineering education improved the results of females, especially.

Due to the results presented in Table 4, it can be concluded that the performance of right-handed students improved significantly by 3.22% among the two tests (p-value = 0.0317). The improvement of results among left-handed students is 0.46%. Thus, engineering education improved mainly the results of right-handed students.

It is shown by the results in Table 6 that while there is no detectable significant difference among the results of all groups, mechatronics students performed significantly better than civil engineering students by 8.83% on the post-test (p-value = 0.041). The improvements between the two tests are 1.18% for civil engineering and 1.79% for mechanical engineering students which are not significant. As was mentioned earlier, mechatronics students did not take the pre-test. This means that when looking from the perspective of different studies, no significant improvements were found among the two tests. It should be noted that however, the students' average rates of correct answers improved between the two tests, although not significantly. This is possibly due to the small timeframe between the two tests.

As can be seen in Table 7, mainly 1st-year students took the tests. According to p-value = 0.045, they improved significantly between the two tests by approximately 2.87%.

According to Table 9, there are differences in the results among various age groups and tests. Those students who are over 18 performed significantly better on the pre-test by approximately 9.96% than those who are 18 or younger (p-value = 0.044). This same group also performed significantly better on the post-test by approximately 13.74% (p-value = 0.019). Also, only this group improved significantly between the tests by approximately 2.86% (p-value = 0.050). The performance of those who are 18 or younger actually decreased by 3.28%. It should be noted that the number of those who are 18 years old at most is quite small on the post-test, and this decrease is fortunately not significant (p-value = 0.570).

Lastly, the test types were assessed. According to Table 11, there are only significant improvements on the MRT test type between the pre-test and post-test by approximately 4.19% (p-value = 0.004). These improvements are only 0.66% and 1.25% in the case of the MCT and PSVT test types, respectively. These two are not significant, but are still improvements, nonetheless. However, this means that engineering education improves the results on the MRT test type, especially.

4.3 Limitations of the Study

Naturally, this study also has its limitations. The main limitation was the motivation, and thus, attitude of the students. Since they knew that the post-test did not affect their grades or did not give them bonus points on exams, some of them just quickly selected random choices on the tests. With proper motivation, students could take the tests more seriously. Another limitation was that the tests were only done using a desktop display. However, conducting the tests on paper or using the Gear VR as well, could complement the results.

Conclusions

As mentioned in the introductory section, since a connection exists between CogInfoCom and Education 4.0, the authors wanted to investigate whether real-life spatial ability training affects this skill in VR. For this goal, the authors used their previously developed spatial ability testing application to measure these skills of 240 university students. A pre-test and a post-test were conducted in the same semester: the former in its beginning and the latter in its end.

According to the results, real-life education has significantly affected the spatial skills in virtual environments. These results, however, vary among different test types and user groups. The users' performance on each test type is improved, although it is only significant in the case of the MRT test type. Also, while the performance of males is significantly better than that of females, it is undeniable that real-life engineering education more strongly affected the performance of females than that of males.

The gap between their performances, became smaller on the post-test, after they attended the spatial ability training courses. Also, those students who were over 18 years of age were significantly affected by these courses. Similarly, right-handed students were more affected by these real-life spatial ability training courses than left-handed ones.

These results show the importance of an engineering education, while they can also strengthen the aspect of a connection exists between CogInfoCom and Education 4.0. Based on the results, it can be concluded that not only the use of VR affects education, but education also influences the use of VR.

Acknowledgement

Project TKP2020-NKA-10 has been implemented with the support provided by the National Research, Development and Innovation Fund of Hungary, financed under the 2020-4.1.1-TKP2020 Thematic Excellence Programme 2020 – National Challenges sub-program funding scheme. The authors would like to thank this support. The authors acknowledge the financial support of Széchenyi 2020 under the EFOP-3.6.1-16-2016-00015 project.

The authors would also like to thank Ms. Mónika Szeles for her help in developing the application as well as Mr. Lóránt Horváth for his help in the creation of the objects for the MCT test type. Last, but not least, the authors would like to thank everyone who helped by testing the spatial ability application.

References

[1] Craig L. Miller: Enhancing visual literacy of engineering students through the use of real and computer generated models. Engineering Design Graphics Journal, Vol. 56, No. 1, pp. 27-38, 1992

[2] Craig L. Miller, Gary R. Bertoline: Spatial visualization research and theories: Their importance in the development of an engineering and technical design graphics curriculum model. Engineering Design Graphics Journal, Vol. 55, No. 3, pp. 5-14, 1991

[3] Steven G. Vandenberg, Allan R. Kuse: Mental rotations, a group test of three-dimensional spatial visualization. Perceptual and motor skills, Vol. 47, No. 2, pp. 599-604, 1978

[4] College Entrance Examination Board (CEEB): Special aptitude test in spatial relations, developed by the college entrance examination board, 1939

[5] Roland B. Guay: Purdue spatial visualization test – visualization of rotations. W. Lafayette, IN. Purdue Research Foundation, 1977

[6] Albert A. Rizzo, J. Galen Buckwalter, Ulrich Neumann, Carl Kesselman, Marcus Thiébaux, Peter Larson, Andre van Rooyen: The virtual reality mental rotation spatial skills project. CyberPsychology & Behavior, Vol. 1, No. 2, pp. 113-119, 1998

[7] Károly Hercegfi, Anita Komlódi, Bálint Szabó, Máté Köles, Emma Lógó, Balázs P. Hámornik, Gyöngyi Rózsa: Experiences of virtual desktop collaboration experiments. In 2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 375-379, 2015

[8] Rafael Molina-Carmona, María Luisa Pertegal-Felices, Antonio Jimeno-Morenilla, Higino Mora-Mora: Virtual reality learning activities for multimedia students to enhance spatial ability. Sustainability, Vol. 10, No. 4, pp. 1074, 2018

[9] Hugo César Gómez-Tone, Jorge Martin-Gutierrez, Lili Valencia Anci, Carlos E. Mora Luis: International Comparative Pilot Study of Spatial Skill Development in Engineering Students through Autonomous Augmented Reality-Based Training. Symmetry, Vol. 12, No. 9, pp. 1401, 2020

[10] Róbert Tóth, Marianna Zichar, Miklós Hoffmann: Gamified Mental Cutting Test for enhancing spatial skills. 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 299-304, 2020

[11] Róbert Tóth, Marianna Zichar, Miklós Hoffmann: Improving and Measuring Spatial Skills with Augmented Reality and Gamification. In International Conference on Geometry and Graphics. Springer, Cham, pp. 755-764, 2021

[12] Péter Baranyi, Ádám Csapó: Definition and synergies of cognitive infocommunications. Acta Polytechnica Hungarica, Vol. 9, No. 1, pp. 67-83, 2012

[13] Gyula Sallai: The cradle of cognitive infocommunications. Acta Polytechnica Hungarica, Vol. 9, No. 1, pp. 171-181, 2012

[14] Peter Baranyi, Adam Csapo, Peter Varlaki: An overview of research trends in coginfocom. In IEEE 18th International Conference on Intelligent Engineering Systems INES 2014, pp. 181-186, 2014

[15] Péter Baranyi, Adam Csapo, Gyula Sallai: Cognitive Infocommunications (CogInfoCom). Springer, 2015

[16] Robert Pinter, Sanja Maravic Cisar: Measuring Team Member Performance in Project Based Learning. Journal of Applied Technical and Educational Sciences, Vol. 8, No. 4, pp. 22-34, 2018

[17] Attila Kovari, Jozsef Katona, Cristina Costescu: Evaluation of Eye-Movement Metrics in a Software Debbuging Task using GP3 Eye Tracker. Acta Polytechnica Hungarica, Vol. 17, No. 2, pp. 57-76, 2020

[18] Attila Kovari, Jozsef Katona, Cristina Costescu: Quantitative Analysis of Relationship Between Visual Attention and Eye-Hand Coordination. Acta Polytechnica Hungarica, Vol. 17, No. 2, pp. 77-95, 2020

[19] Jozsef Katona, Tibor Ujbanyi, Gergely Sziladi, Attila Kovari: Examine the effect of different web-based media on human brain waves. Proceedings of the 8th IEEE International Conference on Cognitive Infocommunications, pp. 407-412, 2017

[20] Bálint Szabó, Károly Hercegfi: Research questions on integrating user experience approaches into software development processes. In 2017 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 243-246, 2017

[21] Jozsef Katona: Clean and Dirty Code Comprehension by Eye-tracking Based Evaluation using GP3 Eye Tracker. Acta Polytechnica Hungarica, Vol. 18, No. 1, pp. 79-99, 2021

[22] Jozsef Katona: Analyse the Readability of LINQ Code using an Eye-Tracking-based Evaluation. Acta Polytechnica Hungarica, Vol. 18, No. 1, pp. 193-215, 2021

[23] Borbála Berki: Experiencing the Sense of Presence within an Educational Desktop Virtual Reality. Acta Polytechnica Hungarica, Vol. 17, No. 2, pp. 255-265, 2020

[24] Tamás Budai, Miklós Kuczmann: Towards a modern, integrated virtual laboratory system. Acta Polytechnica Hungarica, Vol. 15, No. 3, pp. 191-204, 2018

[25] György Molnár, Beáta Orosz: Visual, own device and experience-based educational methods and possibilities in VET. Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications, pp. 517-520, 2019

[26] György Molnár, Katalin Nagy, Zoltán, Balogh: The role and impact of visualization during the processing of educational materials, presentation options in education and in the virtual space, Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications, pp. 533-537, 2019

[27] György Molnár: How digital and virtual life trapped in visuality and new media systems affect teaching and the learning process, Virtual Reality − Real Visuality : Virtual, pp. 101-120, 2017

[28] György Molnár, Dávid Sik: Smart devices, smart environments, smart students – A review on educational opportunities in virtual and augmented reality learning environments, Proceedings of the 10th IEEE International Conference on Cognitive Infocommunications, pp. 495-498, 2019

[29] Robert Demeter, Attila Kovari: Importance of digital simulation in the competence development of engineers defining the society of the future, CIVIL SZEMLE, Vol. 17, No. 2, 2020, pp. 89-101

[30] Miroslav Macik: Cognitive aspects of spatial orientation. Acta Polytechnica Hungarica, Vol. 15, No. 5, pp. 149-167, 2018

[31] Tibor Guzsvinecz, Éva Orbán-Mihálykó, Erika Perge, Cecília Sik-Lányi: Analyzing the Spatial Skills of University Students with a Virtual Reality Application using a Desktop Display and the Gear VR, Acta Polytechnica Hungarica, Vol. 17, No. 2, pp. 35-56, 2020

[32] Tibor Guzsvinecz, Cecilia Sik-Lanyi, Eva Orban-Mihalyko, Erika Perge: The Influence of Display Parameters and Display Devices over Spatial Ability Test Answers in Virtual Reality Environments. Applied Sciences, Vol. 10, No. 2, pp. 526, 2020

[33] Attila Kovari: Study of Algorithmic Problem-Solving and Executive Function, Acta Polytechnica Hungarica, Vol. 17, No. 9, 2020, pp. 241-256

[34] Attila Kovari: Human-machine communication from an industry 4.0 perspective and its relationship with education 4.0, New challenges and pedagogical innovations in VET & HE, 2018, pp. 637-647

[35] Attila Kovari, Jozsef Katona, Ilona Heldal, Robert Demeter, Cristina Costescu, Adrian Rosan, Serge Thill, Teodor Stefanut: Relationship between Education 4.0 and Cognitive InfoCommunications, In 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 431-436, 2020

[36] Andreas Dünser, Karin Steinbügl, Hannes Kaufmann, Judith Glück: Virtual and augmented reality as spatial ability training tools. In Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI, pp. 125-132, 2006

[37] Grigore C. Burdea, Philippe Coiffet: Virtual reality technology. John Wiley & Sons, 2003

[38] Ilona Heldal: The impact of social interaction on usability for distributed virtual environments. International Journal of Virtual Reality, Vol. 6, No. 3, pp. 45-54, 2007

[39] Ralph Schroeder, Ilona Heldal, Jolanda Tromp: The usability of collaborative virtual environments and methods for the analysis of interaction. Presence: Teleoperators and Virtual Environments, Vol. 15, No. 6, pp. 655-667, 2006

[40] Martin Magdin, Zoltán Balogh, Jaroslav Reichel, Jan Francisti, Stefan Koprda, György Molnár: Automatic detection and classification of emotional states in virtual reality and standard environments (LCD): comparing valence and arousal of induced emotions, Virtual Reality, Vol. 25, No. 1, pp. 1-13, 2021