International Education Studies; Vol. 16, No. 4; 2023
ISSN 1913-9020 E-ISSN 1913-9039
Published by Canadian Center of Science and Education

Comparing the Effectiveness of Newer Linework on the Mental Cutting Test (MCT) to Investigate Its Delivery in Online Educational Settings

Theresa Green[1], Wade H. Goodridge[2], Jon Anderson[2], Eric Davishahl[3] & Daniel Kane[2]

[1] School of Engineering Education, Purdue University, West Lafayette, Indiana, USA

[2] Department of Engineering Education, Utah State University, Logan, Utah, USA

[3] Department of Sciences, Whatcom Community College, Bellingham, Washington, USA

Correspondence: Wade H. Goodridge, Department of Engineering Education, Utah State University, Logan, UT., 84321, USA. Tel: 1-435-797-9051. E-mail: wade.goodridge@usu.edu

Received: December 21, 2022 Accepted: April 4, 2023 Online Published: July 16, 2023

doi:10.5539/ies.v16n4p1 URL: https://doi.org/10.5539/ies.v16n4p1

Abstract

The purpose of this study was to examine any differences in test scores between three different online versions of the Mental Cutting Test (MCT). The MCT was developed to quantify a rotational and proportion construct of spatial ability and has been used extensively to assess spatial ability. This test was developed in 1938 as a paper-and-pencil test, where examinees are presented with a two-dimensional drawing of a 3D object containing a cutting plane passing through the object. The examinee must then determine the cross-sectional shape that would result from cutting along the imaginary cutting plane. This work explored three versions of this test (the original and two adapted versions), administered online, to see if there were any differences on the versions regarding student performance. Versions differed in the linework quality displayed as well as shading shown on the surfaces.

This study analyzed statics students' scores on the three online versions of the MCT and on the original paper version of the MCT to identify which version of the test may be most optimal for administering to engineering students. Results showed that there was a statistically significant difference in students' scores between multiple versions. Understanding which representations of the MCT items are most clear to students will provide insights for educators looking to improve and understand the spatial ability of their students.

Keywords: engineering education, spatial education, mental cutting test, MCT, online assessment, spatial ability, spatial cognition

1. Introduction and Background

1.1 Spatial Ability

Spatial ability can be defined as one's ability to mentally represent, construct, or transform information about objects in physical space (Linn & Petersen, 1985; Newcombe & Shipley, 2015). Spatial ability and spatial skills are essential in most people's daily lives in order to navigate the world and environment, interact with objects, and use tools and physical manipulatives. There are different ways that spatial thinking can be operationalized for investigation. Linn and Petersen (1985) identified three categories of spatial thinking that are commonly explored in spatial ability research: Visualization, mental rotation, and spatial perception.

Spatial visualization is defined by Linn and Petersen (1985) as "tasks that involve complicated, multistep manipulations of spatially presented information" (p. 1484). The authors present examples of tasks that utilize spatial visualization skills, such as mental paper folding, a test where someone is presented with a two-dimensional figure and must determine what the resulting shape would look like if it were folded a certain way. *Mental rotation* is the ability for someone to rotate a two or three-dimensional shape in their mind (Linn & Petersen, 1985). One test that has been developed to measure mental rotation is the Mental Rotation Test developed by Vandenberg and Kuse (1978). In this test, participants view a drawing of a three-dimensional figure and must choose, from the four given responses, the two that correctly show the drawing in a rotated view. *Spatial perception* is the ability to interpret spatial relationships between oneself and other objects (Linn &

Petersen, 1985). Linn and Petersen (1985) describe how one method for measuring this construct is through the water-level task, where participants must identify a horizontal line that appears in a tilted glass.

Research in spatial ability has suggested that spatial skills can be trained using targeted interventions (Uttal et al., 2013). In their meta-analysis on the effects of spatial training, Uttal and colleagues found that spatial skills are malleable and can be trained using a variety of interventions, such as playing video games or taking courses that required the use of spatial skills (e.g., engineering graphics) (Uttal et al., 2013). Further work has found gains in spatial scores when students complete tasks such as manually sketching three-dimensional objects on a two-dimensional plane, sketching 3D objects given a 2D drawing, and manipulating virtual 3D objects using augmented reality (Martin-Gutiérrez & González, 2016, pp. 225-239) In addition, Uttal et al.'s (2013) meta-analysis revealed that spatial skills are durable and transferable. This suggests that spatial ability skills can persist over time, even months after a spatial intervention, and that these skills can transfer to other situations are related to those conducted in the spatial trainings.

Several instruments have been developed to measure various constructs of spatial ability. Some of these instruments include the Mental Cutting Test (MCT), the Purdue Spatial Visualization Test: Visualization of Rotations (PSVT:R), and the Mental Rotation Test (MRT). The MCT was developed by the College Entrance Examination Board (CEEB) in 1938 as part of a college admittance examination (CEEB, 1939). The MCT has been used by various researchers (e.g., Gorska & Sorby, 2008; Wood, Goodridge, Call, & Sweeten, 2016) to explore different elements of spatial ability including spatial visualization and spatial relations. Wood et al. (2016) have pointed out that it may also address concepts of proportion. The test consists of 25 items each with five different answer choices, intended to be completed in 20 minutes. For each test item, test-takers are presented with a two-dimensional drawing of a three-dimensional object that contains a cutting plane that passes through a portion of the object. Each item then has five different answer choices from which the test-taker must choose the response that most closely resembles what the cross-sectional area of the surface made from the cutting plane would look like. Figure 1 shows an example of an item on the MCT that would be shown to test-takers.

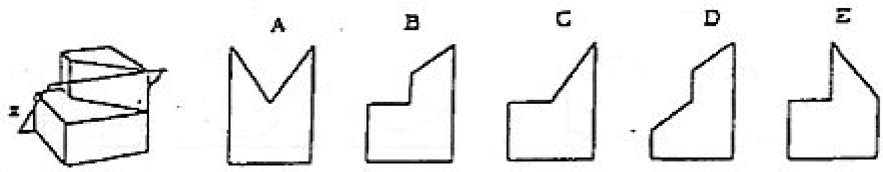

Figure 1. An example test item from the Mental Cutting Test (MCT). Test-takers are presented with a 2D drawing of a 3D object with a cutting plane passing through it and five answer choices representing the cross-section made by the cutting plane

The PSVT:R is a test that was developed by Guay in 1977 to measure mental rotation ability (Guay, 1977). This test consists of 30 multiple choice items, where each item has five possible answer choices. On each question, test-takers are presented with an object in one orientation followed by the same object after being rotated in some manner. A second object is also shown, and test-takers must select, from the five answer choices, the correct depiction of the second object after it has been rotated in the same manner as the first object. A sample question from the PSVT:R is shown in Figure 2.

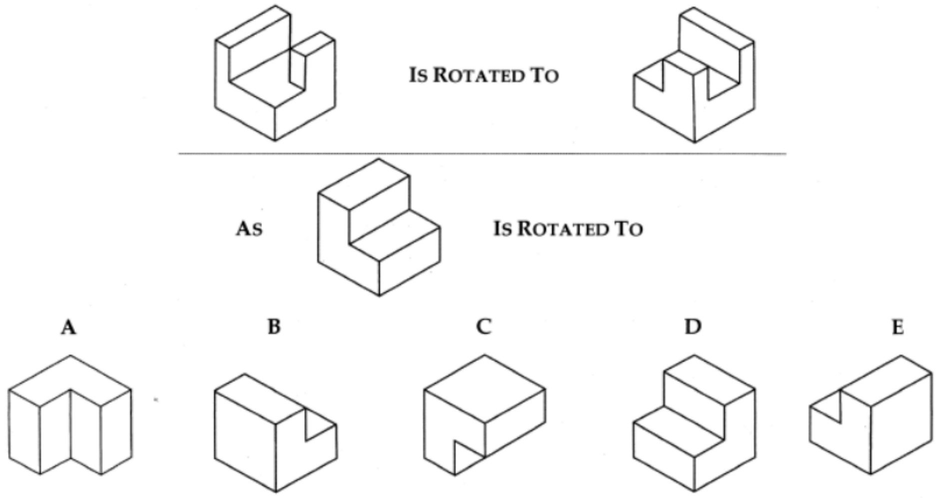

Figure 2. Example question from the Purdue Spatial Visualization Test: Visualization of Rotations (PSVT:R) (Guay, 1977). Based on the rotation of the first object, test-takers must choose the answer corresponding to the correct rotation of the second object

The MRT was developed by Vandenberg and Kuse in 1978 and assesses the test-taker's mental rotation skills (Vandenberg & Kuse, 1978). The MRT contains 20 test items, each with four answer choices. Each item contains a criterion figure, a two-dimensional representation of a three-dimensional object. The examinee must choose, from the four responses, the two answer choices that show correctly rotated representations of the criterion figure. The other two answer choices are incorrect, distractor responses. An example problem from the MRT is presented in Figure 3.

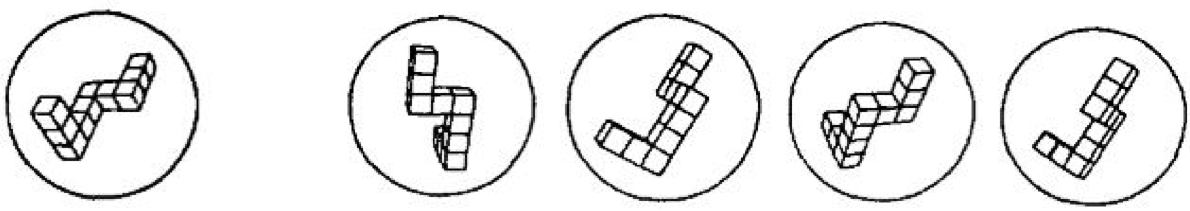

Figure 3. Example problem from the Mental Rotation Test (MRT) (Vandenberg & Kuse, 1978). Test-takers must choose the two responses from the four answer choices that show a correct rotation of the criterion figure

1.2 Spatial Ability in Engineering

Spatial skills are important for success in science, technology, engineering, and mathematics (STEM) fields, including biology, geology, chemistry, mathematics, and engineering (Hegarty, 2014). For example, in chemistry, it is important to be able to understand spatial relationships between three-dimensional molecules and configurations of different types of molecules (Uttal & Cohen, 2012), and interacting with three dimensional molecule models can increase spatial thinking (Stull & Hegarty, 2016). In geoscience fields, spatial skills are important when reading and interpreting maps or visualizing objects and scenery from different points and perspectives (Newcombe & Shipley, 2015). In addition, spatial ability has been shown to significantly predict success in STEM fields (Uttal & Cohen, 2012; Wai et al., 2009). In their study, Wai et al. (2009) used longitudinal data taken from students 11 years after they graduated from high school to identify how spatial ability impacted their future occupations or educational pursuits. This effort found that spatial ability was an important skill for predicting success in STEM careers and achieving advanced STEM degrees.

In engineering specifically, spatial skills such as spatial visualization and mental rotation have been shown to be important (Hsi, Linn, & Bell, 1997). Spatial skills allow students to mentally visualize word problems and translate problem statements into mathematical equations (Duffy, Sorby, & Bowe, 2020). These skills are also necessary for understanding and interpreting engineering graphics (Sorby & Baartmans, 2000) and for computer programming (Román-González et al., 2017). In their study of the impacts of including spatial instruction in an engineering course, Hsi et al. (1997) found that an intervention to target spatial skill development improved students' course grade and their spatial reasoning skills. In addition, in their interviews with practicing engineers, the authors found that engineering professionals perceived spatial skills to be important for engineering practice and that they used a variety of spatial reasoning strategies to solve problems.

The MCT, PSVT:R, and MRT have all been used by various researchers to understand the spatial ability skills of engineering students. Since 1996, Sorby and Baartmans (2000) have used the MCT and the MRT as pre- and post-tests in their preparatory spatial visualization course at Michigan Technological University (MTU) to help support students' spatial skill development before taking engineering graphics courses. Sorby et al. (2020) also used the MCT to measure chemical engineering students' spatial ability to understand the relationship between spatial skills and problem solving and between spatial ability and gender. Leopold (2005) used the MCT and the MRT to test undergraduate civil and architectural engineering students' spatial ability at the beginning of their educational careers. In addition, the PSVT:R has been used at MTU as a pre- and post-test to understand the spatial ability of engineering students before and after completing a spatial visualization course (e.g., Gorska & Sorby, 2008). The use of all three of these instruments has led to various insights about engineering students' spatial ability skills and the impacts these skills have on performance in engineering subjects. The present study focused on the MCT specifically with an eye to its administration through an LMS as well as a focus on improving its linework and implementing shading.

2. The Present Study: Adaptation of the MCT

One goal of this study was to explore the viability of implementing the MCT in an online format. Traditionally, the MCT is offered as a paper-and-pencil test. In an effort to make the MCT more accessible in large classrooms and for easier administration, the test was adapted by the research team for electronic administration through learning management systems online, such as Canvas.

The second goal of this study was to explore the effect of updating the problem representations on the MCT. To explore this idea, the research team developed two adapted versions of the MCT, each with updated 2D drawings and linework. The first adapted version of the MCT was developed in 2019. This version included updated 2D drawings of each of the 3D objects that were presented in the original MCT. The updated drawings were developed in a computer-aided design (CAD) software program with the intention of improving the clarity and size of the images. The drawings created in the CAD program were larger than the original images on the 1938 version of the MCT. In addition, the cutting plane passing through the 3D object in the drawings was displayed as translucent in a different color shading than the color of the 3D object. The cutting plane shading stopped where the cutting plane intersected the object. In the original 1938 version of the MCT, the plane was drawn completely transparent and there was no shading on it whatsoever. Linework on the answer choices was also improved in CAD and were also shaded in color on the 2019 test version.

The second adaptation of the MCT was developed in 2021. Similar to the 2019 version, the 2021 version was developed to improve the clarity and size of the 2D images presented in the test. The items in this version of the test were also created in a CAD program, but the enlarged images also included color shading on the 2D drawings of the 3D objects. The shading was on the surfaces of the object and was darker and lighter to reveal a direction of lighting. In addition, the cutting plane was again completely transparent with no shading, similar to the 1938 version of the MCT. The elimination of the cutting plane shading was motivated in part by an observation that the shading in the 2019 version yielded higher student scores, likely because the shading contributed to a clearer depiction of the cross-sectional shape. Similar to the 2019 version, the linework on the answer choices was improved in CAD but they also did not include any shading.

In both the 2019 and 2021 adaptations of the MCT, no new test items were added. Each question from the 1938 version of the MCT was adapted with updated drawing representations for each new test version. Figure 4 presents an example problem from the MCT with a side-by-side comparison of drawing representations of the item and its potential answer choices across the three test versions.

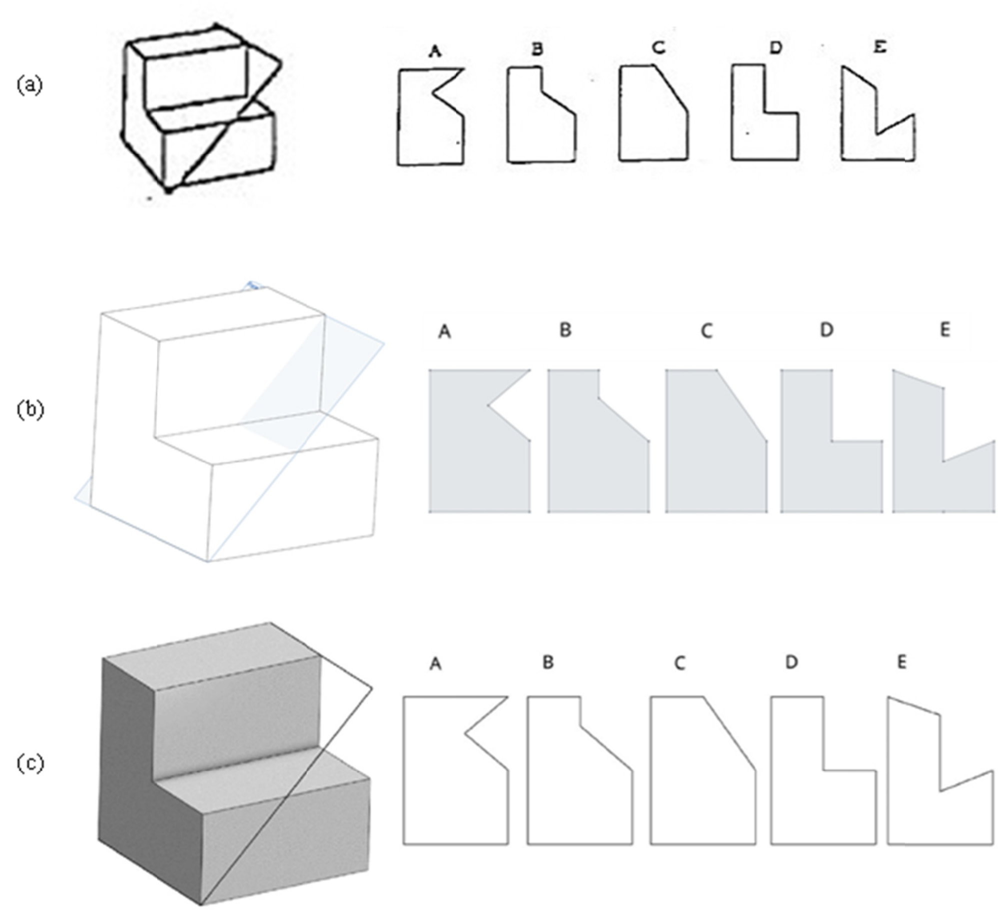

Figure 4. An example problem from the Mental Cutting Test (MCT) with the corresponding answer choices for the (a) 1938 version, (b) the 2019 version, and (c) the 2021 version

The purpose of this paper is to explore differences in students' scores and item difficulty on these three online versions of the MCT and to make comparisons to scores and item difficulty on the original 1938 test version administered on paper. In addition, this study aimed to identify which online test version may be best for educators to implement in their own courses to measure students' spatial ability.

3. Methods

3.1 Participants and Procedure

Test score data from 174 students enrolled in the fall 2021 semester of an engineering statics course at a large Western university was collected for this study. An institutional review board (IRB) protocol was developed at the research team's university to ensure the ethical handling of student test score data. Students in the fall 2021 offering of the course were randomly assigned to take one of the three versions of the MCT as an opportunity to earn extra credit in the statics course. We used Canvas, the Learning Management System (LMS) used at our university, to sort students randomly into which version of the exam they would be tested with. The group feature in Canvas allows for the creation of different groups within a group category, and one of the features inside of Canvas is the assignment of students to the different groups at random. Once the groups are created and the students assigned, then each group received a different version of the MCT quiz by assigning the quiz to only that group. Students were assigned to groups in Canvas randomly by the LMS system. In this way, students were only able to access their own assigned quizzes and did not know what the other quizzes were. Students took these tests near the end of the semester. Each version of the MCT was converted into a multiple-choice quiz on Canvas and was included as an extra credit assignment on the statics course page. Students were limited to 20 minutes to take the test. In total, 59 students took the 1938 version of the MCT, 54 students took the 2019 version of the MCT, and 58 students took the 2021 version of the MCT.

A paper version of the 1938 version of the MCT was also administered to statics students in the spring 2022 semester at the same university under the same IRB protocol. A total of 54 students took this version of the MCT. This test was offered as an extra credit opportunity and was administered during class time on one day near the end of the spring 2022 semester. As was done for the online versions of the test, students were limited to 20 minutes to take the paper version of the MCT.

3.2 Data Analysis

A one-way analysis of variance (ANOVA) test was performed on the data to compare scores that students earned on the three different versions of the MCT (i.e., the 1938, 2019, and 2021 versions). Each of the three versions of the test was considered as a group for the ANOVA procedure. The analysis was performed in MATLAB version R2022a (MathWorks, 2022).

In addition, difficulty and discrimination indices were calculated for each question on each of the three test versions. To find these values, the students taking each test were divided into high performers, performers, and low performers. The percentage of correctly responding students to the questions was calculated for the high performers and the low performers for each question. The item difficulty is calculated as the average percentage of high performer and low performer percentage on each question. The item difficulty is found by subtracting the low performer percentage from the high performer percentage.

Last, a Chi-Squared Test of Independence was performed to assess the relationship between test version and whether students answered each of the 25 questions correctly.

4. Results

4.1 Internal Reliability of MCT Tests

Mcdonald's omega and Cronbach's alpha were determined for each exam separately to determine the internal reliability of each test (Dunn et al, 2014; Flora, 2020). The results of each test are shown in Table 1. Because one question in the 2019 test received a correct answer by all participants and thus had no standard deviation, R could not find either the omega or alpha. To alleviate this, that question was removed from the reliability measure. All three tests fall within accepted parameters (ω_t and α are greater than .70), although the 2019 MCT test showed the highest reliability.

Table 1. Internal reliability for each MCT Test

MCT Test	Items in Analysis	Number of responses	ω_t	α
1938	26	59	.79	.75
2019	25	54	.92	.90
2021	26	58	.86	.83

4.2 ANOVA Results

The mean test scores for the 1938 online, 2019, and 2021 versions of the MCT were 13.783 (SD = 4.165), 20.400 (SD = 4.604), and 15.458 (SD = 4.931), respectively. The one-way ANOVA revealed that there was a statistically significant difference in the mean MCT score between at least two versions of the MCT ($F(2, 168) = 32.53$, $p < .001$). Tukey's HSD test for multiple comparisons between the groups revealed that the mean test score was significantly different between the 1938 online and 2019 versions of the test ($p < .001$) and was significantly different between the 2021 and 2019 versions of the test ($p < .001$). There was no statistically significant difference in the mean test score between the 1938 online and 2021 versions of the test ($p = .110$). These results are summarized in Table 2.

Table 2. One-way ANOVA results to compare the mean test score on the 1938, 2019, and 2021 versions of the MCT

Predictor	Sum of squares	df	Mean square	F	p
Groups	1361.48	2	680.739	32.53	<.001
Error	3515.9	168	20.928		
Total	4877.38	170			

4.3 Item Analysis Results

Item analysis of the questions on each version of the test was also performed. The results of this analysis is shown in Table 3. The difficulty index and item discriminating power are presented for each problem on each test version. In addition, item analysis was then performed on the paper version of the 1938 MCT based on students' scores. The results of this analysis are presented in Table 4. The results for the online version of the 1938 test are also presented in Table 4 to compare the difficulty and discriminating power between these two versions of the test.

Table 3. Item difficulty and item discriminating power on the 1938, 2019, and 2021 versions of the MCT

		Question number												
	Test version	1	2	3	4	5	6	7	8	9	10	11	12	13
Difficulty	1938	.85	.65	.88	.58%	.80	.80	.63	.88	.40	.50	.83	.50	.30
	2019	.86	.69	1.00	.75	.97	.94	.81	.86	.67	.89	.83	.89	.53
	2021	.71	.63	.95	.79	.82	.79	.76	.74	.58	.47	.71	.66	.61
Discrimination	1938	.20	.70	.25	.25	.20	.10	.45	.25	.00	.60	.25	.40	.50
	2019	.28	.50	.00	.17	.06	.11	.39	.28	.56	.22	.33	.22	.50
	2021	.58	.74	.11	.42	.26	.32	.37	.53	.42	.42	.37	.58	.58
		14	15	16	17	18	19	20	21	22	23	24	25	
Difficulty	1938	.48	.43	.63	.38	.65	.43	.48	.35	.43	.38	.38	.18	
	2019	.92	.72	.92	.86	.75	.78	.75	.81	.86	.67	.58	.42	
	2021	.55	.66	.71	.08	.55	.42	.71	.61	.58	.42	.32	.42	
Discrimination	1938	.45	.65	.45	.35	.40	.15	.55	.40	.55	-.05	.35	.35	
	2019	.17	.56	.17	.28	.39	.44	.50	.28	.28	.56	.72	.61	
	2021	.89	.68	.47	-.16	.79	.53	.37	.37	.53	.42	.32	.42	

Table 4. Item difficulty and item discriminating power on the 1938 version of the MCT administered on paper and online

		Question number												
	Test version	1	2	3	4	5	6	7	8	9	10	11	12	13
Difficulty	1938 Paper	.97	.67	1.00	.58	.86	.78	.81	.83	.67	.69	.89	.67	.39
	1938 Online	.85	.65	.88	.58	.80	.80	.63	.88	.40	.50	.83	.50	.30
Discrimination	1938 Paper	.06	.56	.00	.61	.28	.44	.39	.33	.44	.61	.11	.44	.67
	1938 Online	.20	.70	.25	.25	.20	.10	.45	.25	.00	.60	.25	.40	.50
		14	15	16	17	18	19	20	21	22	23	24	25	
Difficulty	1938 Paper	.64	.61	.78	.42	.67	.67	.81	.42	.56	.50	.44	.22	
	1938 Online	.48	.43	.63	.38	.65	.43	.48	.35	.43	.38	.38	.18	
Discrimination	1938 Paper	.61	.44	.22	.72	.56	.56	.39	.61	.33	.44	.67	.44	
	1938 Online	.45	.65	.45	.35	.40	.15	.55	.40	.55	-.05	.35	.35	

4.4 Chi-Squared Test Results

The results of the Chi-squared test revealed that there was a significant relationship between the two variables, $X^2(72, N = 225) = 99.154$, $p = 0.019$. Cramer's V was also calculated to determine the effect size of the Chi-squared measurement. Cramer's V was calculated to be 0.11, which indicates a small effect. Table 5 shows the results of the Chi-squared test for each individual question across the four test versions.

Table 5. Chi-squared results from analysis of the four versions of the MCT, broken down by the percent contribution of each test question to the overall Chi-squared statistic (99.154)

	Question number												
Test version	1	2	3	4	5	6	7	8	9	10	11	12	13
1938 Paper	0.369	0.001	0.022	0.169	0.050	0.153	0.254	0.025	0.121	0.008	0.380	0.021	0.763
1938 Online	1.169	1.336	0.890	0.106	0.948	0.397	0.011	2.248	0.343	0.030	1.454	0.306	1.190
2019	1.629	1.284	1.878	0.476	0.745	0.296	0.807	1.365	0.112	1.056	1.504	0.219	0.048
2021	0.048	0.034	0.227	2.219	0.066	0.159	0.157	0.001	0.027	1.146	0.161	0.019	4.652
	14	15	16	17	18	19	20	21	22	23	24	25	
1938 Paper	0.059	0.000	0.035	0.021	0.020	0.006	1.648	0.986	0.277	0.292	0.091	0.987	
1938 Online	0.253	1.989	0.022	0.041	0.653	0.397	0.306	2.475	1.503	0.194	0.180	2.129	
2019	0.969	0.149	0.229	14.626	0.009	0.467	1.674	1.012	1.775	1.304	1.159	0.438	
2021	0.139	0.835	0.334	17.684	0.653	0.058	0.458	1.887	0.045	0.096	3.621	2.716	

5. Discussion

5.1 Impact of Shading the Cutting Plane

Based on the average scores on the three versions of the MCT, students scored the highest on the 2019 version of the test as expected, and students scored lowest on the 1938 version of the test. The ANOVA revealed that the scores between the 1938 and 2019 versions of the MCT were statistically significantly different. The new linework and improved representations of the MCT items on the 2019 test version may be one reason why there is improvement in scores from the 1938 version. However, it is more likely that the greatest difference is seen between the 1938 version and the 2019 version due to the translucent shading on the cutting plane which may more readily reveal the cutting plane contact points to the exterior of the original object. Those contact points are distinct as the cutting plane is not shaded after it enters the object (see Figure 4b) and thus an astute observer can begin to discern a three-dimensional perspective of the outline formed from the intersection of the cutting plane to the object, which is observable only from the front viewing perspective. This outline is likely a distinct advantage in discerning an appropriate solution. In the 2021 version, this shading on the cutting plane was nonexistent, creating an image that is more similar to the original 1938 version (except in surface shading on the object; see Figure 4c). The lack of significant difference between scores on the 1938 version and the 2021 version, despite the latter's surface shading, indicates that the shading of the object's surfaces did not contribute to an increase in spatial ability.

Next, the ANOVA revealed that the scores on the 2019 and 2021 test versions were significantly different. This result may also be due to the improved problem representation on the MCT items between these two test versions. The shaded cutting plane on the 2019 version again may have helped test-takers in discerning the correct answer compared to the transparent cutting plane in the 2021 version. The present analysis does not suggest that shading the object as done in the 2021 test version leads to improved outcomes for test-takers. Future work may be needed to uncover the reasons behind this argument. For example, the shaded object may make the problem more difficult to interpret by introducing additional visual elements to the 3D object that make it harder to mentally trace the intersection of the cutting plane.

In addition, future work may need to be conducted to continue to investigate the reasons that there was no significant difference between scores on the 1938 and 2021 test versions. This may be due to the similarity in the transparent representation of the cutting plane that was present in both these versions of the test. This finding also supports the argument that shading the 3D object, as done in the 2021 test version, does not lead to improved scores on the MCT.

The Chi-squared test yielded similar results. In particular, the 2019 and 2021 representations of question 17 contributed 32.31% to the total Chi-square score, and thus account for most of the difference between expected and observed values, suggesting that the shading of the plane or object significantly contributed to students' abilities to answer the question. Similarly, the 2021 representation of questions 4, 10, 13, 21, 24, and 25 contributed a further 24.26% of the total Chi-square score, suggesting that these problems were easier for this version of the test than the other versions.

5.2 Comparison to Paper Version

Mean scores on the 1938 version of the MCT that was administered on Canvas were also compared to the mean scores on the same test version administered on paper. The mean score on the paper version of this test was

16.537 (SD = 4.936).

Previous research by Wood et al. (2016) also explored statics students' mean scores on the 1938 version of the MCT in an online format administered in a formal testing center on a university campus. Their analysis showed that mean scores on this test taken at the end of the semester was 16.537. This mean is similar to the mean obtained from administering the paper version of the 1938 MCT to a different set of statics students. Both means differ from the mean score on the 1938 MCT that was given through Canvas during the fall of 2021. This difference is statistically significant ($t(111) = -3.436$, $p < .001$) and indicates an area for further study. In the study by Wood et al. (2016), the 1938 version of the MCT was also administered online, similar to how the 1938 version of the MCT was given in the present study. However, the mean score on this test version was lower than the mean score in Wood et al.'s (2016) study. Future work may be needed to uncover why these mean scores were statistically significantly different from each other.

6. Limitations

This study is limited by the lack of demographic variables that were able to be captured from the sample populations. While this work followed a more experimental focused method where a random assignment of participants was possible to the three online versions of the instrument, the ability to track demographic constructs of gender, age, etc., was lost in doing so. Thus, these findings are not parsed out based on these demographic variables and future work should look to account for them.

7. Conclusions

By understanding whether versions of the MCT with updated graphics result in improved outcomes for students, adapted versions of the MCT can be more frequently administered. Supporting students in developing spatial ability skills may improve the likelihood of their success and continued persistence in engineering disciplines.

In addition, this study provides evidence that the administration of such spatial tests is practical through online LMS systems. Researchers desiring to use the MCT in an online format can be assured that they are using a viable instrument. This work also indicates that care must be taken in cutting plane shading when and if a researcher decides to adopt a similar graphical evolution on their spatial test items as was done on the 2019 version of the MCT used in this work. We also point out that the shading of objects' surfaces and enlargement of object drawings do not significantly impact results from the original 1938 version of the MCT.

8. Future Work

Future work should extend towards exploring the lack of significant differences in the means of the 1938 online and 2021 versions of the MCT. In addition, future work should investigate the reasons behind why the mean score on the paper 1938 version of the MCT was higher than the online version of the same test. Last, future work should be conducted to supplement these results with a larger population of students with expanded demographics. The online versions of this instrument, deliverable through a LMS system, can easily facilitate this type of work for researchers at other institutions.

References

College Entrance Examination Board. (1939). *CEEB Special Aptitude Test in Spatial Relations*.

Duffy, G., Sorby, S., & Bowe, B. (2020). An investigation of the role of spatial ability in representing and solving word problems among engineering students. *Journal of Engineering Education, 109*(3), 424-442. https://doi.org/10.1002/jee.20349

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British journal of psychology, 105*(3), 399-412. https://doi.org/10.1111/bjop.12046

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science, 3*(4), 484-501. https://doi.org/10.1177/2515245920951747

Gorska, R., & Sorby, S. (2008). Testing instruments for the assessment of 3D spatial skills. *2008 American Society for Engineering Education (ASEE) Annual Conference & Exposition*. https://doi.org/10.18260/1-2--4411

Guay, R. B. (1977). *Purdue spatial visualization test: Visualization of rotations*. Purdue Research Foundation.

Hegarty, M. (2014). Spatial Thinking in Undergraduate Science Education. *Spatial Cognition & Computation, 14*(2), 142-167. https://doi.org/10.1080/13875868.2014.889696

Hsi, S., Linn, M. C., & Bell, J. E. (1997). Role of spatial reasoning in engineering and the design of spatial instruction. *Journal of Engineering Education, 86*(2), 151-158. https://doi.org/10.1002/j.2168-9830.1997.tb00278.x

Leopold, C. (2005). Geometry education for developing spatial visualisation abilities of engineering students. *The Journal of Polish Society for Geometry and Engineering Graphics, 15*, 39-45.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development, 56*(6), 1479. https://doi.org/10.2307/1130467

Martin-Gutiérrez, J., & González, M. M. A. (2016). Ranking and Predicting Results for Different Training Activities to Develop Spatial Abilities. In M. S. Khine (Ed.), Visual-spatial ability in STEM education (pp. 225-239). Springer. https://doi.org/10.1007/978-3-319-44385-0_11

MathWorks. (2022). *MATLAB* (No. R2022a). Retrieved from https://www.mathworks.com/products/matlab.html?s_tid=hp_ff_p_matlab

Newcombe, N., & Shipley, T. F. (2015). Thinking about spatial thinking: New typology, new assessments. In J. S. Gero (Ed.), *Studying Visual and Spatial Reasoning for Design Creativity*. Springer. https://doi.org/10.1007/978-94-017-9297-4

Román-González, M., Pérez-González, J.-C., & Jiménez-Fernández, C. (2017). Which cognitive abilities underlie computational thinking? *Criterion validity of the Computational Thinking Test. Computers in Human Behavior, 72*, 678-691. https://doi.org/10.1016/j.chb.2016.08.047

Sorby, S. A., & Baartmans, B. J. (2000). The development and assessment of a course for enhancing the 3-D spatial visualization skills of first year engineering students. *Journal of Engineering Education, 89*(3), 301-307. https://doi.org/10.1002/j.2168-9830.2000.tb00529.x

Sorby, S. A., Duffy, G., & Loney, N. (2020). An examination of the role of spatial ability in the process of problem solving in chemical engineering. *Australasian Journal of Engineering Education, 25*(1), 55-65. https://doi.org/10.1080/22054952.2020.1785653

Stull, A. T., & Hegarty, M. (2016). Model manipulation and learning: Fostering representational competence with virtual and concrete models. *Journal of Educational Psychology, 108*, 509-527. https://doi.org/10.1037/edu0000077

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how? In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (pp. 147-181). Elsevier Inc. https://doi.org/10.1016/B978-0-12-394293-7.00004-2

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin, 139*(2), 352-402. https://doi.org/10.1037/a0028446

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*(2), 599-604. https://doi.org/10.2466/pms.1978.47.2.599

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817-835. https://doi.org/10.1037/a0016127

Wood, S. D., Goodridge, W. H., Call, B. J., & Sweeten, T. L. (2016). Preliminary analysis of spatial ability improvement within an engineering mechanics course: Statics. *2016 American Society for Engineering Education (ASEE) Annual Conference & Exposition*. https://doi.org/10.18260/p.25942

Copyrights