Performance by Gender on University Placement Tests in Mathematics and Spatial Skills

Mr. Gavin Duffy, Ohio State University

For the current academic year I am a visiting scholar at Ohio State University with my home institute being Dublin Institute of Technology where I am a lecturer in the School of Electrical & Electronic Engineering. AT OSU I am working on a research project that is investigating the relationship between spatial ability and problem solving in engineering education. Before joining academia I worked in industry as a chemical engineer and control systems engineer and those are the topics that I teach in the DIT.

Dr. Sheryl A. Sorby, Ohio State University

Dr. Sheryl Sorby is currently a Professor of STEM Education at The Ohio State University and was recently a Fulbright Scholar at the Dublin Institute of Technology in Dublin, Ireland. She is a professor emerita of Mechanical Engineering-Engineering Mechanics at Michigan Technological University and the P.I. or co-P.I. on more than $9M in grant funding, most for educational projects. She is the former Associate Dean for Academic Programs in the College of Engineering at Michigan Tech and she served at the National Science Foundataion as a Program Director in the Division of Undergraduate Education from January 2007 through August 2009. Prior to her appointment as Associate Dean, Dr. Sorby served as chair of the Engineering Fundamentals Department at Michigan Tech. She received a B.S. in Civil Engineering, an M.S. in Engineering Mechanics, and a Ph.D. in Mechanical Engineering-Engineering Mechanics, all from Michigan Tech. Dr. Sorby has a well-established research program in spatial visualization and is actively involved in the development of various educational programs.

Austin Mack, Ohio State University
Prof. Brian Bowe, Dublin Institute of Technology

Brian Bowe is the Head of Learning Development in the College of Engineering & Built Environment, Dublin Institute of Technology (DIT). Brian holds a BSc in Applied Science (Physics & Mathematics), a MA in Higher Education and a PhD (Physics) from Trinity College Dublin. In 2000, as an academic member of the School of Physics DIT, he formed the Physics Education Research Group. In 2008, he established an Engineering Education Research Group and in September 2013, he, along with colleagues, created a new research group to encompass all education research activities in the College: "Contributions to Research in Engineering, Architecture & Technology Education" (CREATE). This new education research group already has 18 faculty members, 6 PhD students and a visiting professor from the US. Brian has supervised 8 postgraduate students to completion (5 PhDs & 3 MPhils) and is currently supervising 5 PhD students, all engaged in education research. Brian has also facilitated over 300 education development workshops on problem-based learning, assessment, curriculum development and peer instruction across 10 countries. His education research interests include examining students' approaches to learning within group-based project-driven pedagogies, epistemological development, progression, conceptual understanding and pedagogical evaluations.

Performance by gender on university placement tests in mathematics and spatial skills

Introduction

In an effort to enhance the first year experience (FYE) it is now common for higher education institutes to coordinate orientation programmes for incoming freshman students. Reaching out to students in advance of their arrival at college is a good example of the '*proactive management of student transition*' recommended by Yorke & Longden (2008) who conducted a large study of retention in higher education in the UK. Aptitude and core competency testing is often an integral part of these orientation programmes and while it has been common in the context of engineering education for such testing to be limited to mathematics, the current trend is to also include spatial ability tests on the orientation agenda. The reason is simple: success in STEM education is better predicted by a combination of math and spatial ability scores than by just math alone (Shea, Lubinski, & Benbow, 2001).

Project Talent, undertaken in the US in the 1960s, involved the administration of a battery of psychometric tests over a one week period to a very large sample of high school students. 50,000 males and 50,000 females were recruited from each of grades 9 to 12 (i.e. total n = 400,000) to participate in the study and they were tracked over time (1, 5 and 11 years after the initial tests) to determine whether or not they pursued higher education and, if so, what courses they selected and the highest level of qualification they achieved. Results showed a marked difference in the verbal/spatial/mathematical ability profiles (as measured in high school) of those who were destined to pursue a humanities social science (HSS) education versus those headed for STEM education (Wai, Lubinski, & Benbow, 2009). The latter group had, on average, a much higher spatial ability score than the former. Likewise for mathematical ability. Within the STEM group, spatial ability increased with highest level of education achieved, i.e. spatial aptitude scores measured in high school were much higher for those who were to gain STEM qualifications at PhD level compared to Bachelors level.

The majority, 90 %, of STEM PhD graduates from this sample could be traced back to the group in the top quartile in spatial ability in high school. The results from this analysis are shown in Figure 1 as triads of average scores separated by eventual profession. The score on the left is verbal (V) ability, in the middle is spatial (S) and math (M) is on the right. STEM students, to the right of Figure 1, have an 'I' shaped ability profile (i.e. M > S > V), in contrast to the 'V' shaped profile (i.e. M ≥ V > S) of the HSS students. Clearly, the 'I' shaped profile, developed by high school, was a predictor of a STEM education path and distance travelled on this path. Given that this predictor contains not just math ability but spatial ability also, STEM educators have reason to treat spatial ability in the same way as math ability: assess incoming students for the ability and provide resources to address any shortcomings in it. While it is now common to find math learning support centers co-existing beside engineering schools, it is unusual to find resources made available to improve spatial skills. The findings of Wai et al. (2009) raise spatial skills development as a potentially fruitful way to make STEM education and careers more attractive and to improve grades and retention rates in engineering education.

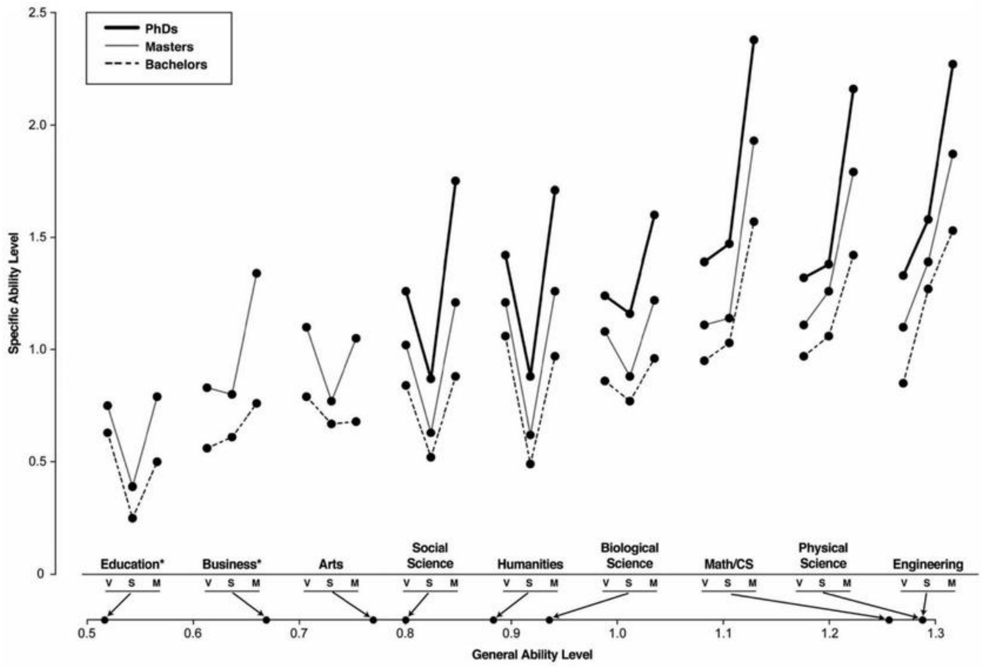


Figure 1. Analysis of Project TALENT data to show relative position of spatial scores to verbal and math scores for different disciplines; V = Verbal, S = Spatial and M = Mathematical ability; (Figure B1 taken from (Wai et al., 2009)).

One of the most interesting findings from spatial ability research, and which is also very important for engineering educators to be aware of, is the sizeable and significant gender gap in favour of males – on average, males get higher scores than females on measures of spatial ability and this has been shown to be the case across the globe (Lippa, Collaer, & Peters, 2010). Women are, therefore, more likely than men to have the 'V' shaped profile thereby placing them at a cognitive disadvantage when studying STEM subjects. Furthermore, if those measured as having a 'V' shaped profile (male and female) were to improve spatial ability and move towards the 'I' shaped profile then they would increase the probability of success in STEM education.

Uttal et al. (2013) reviewed several studies that reported improvements in spatial ability due to spatial skills intervention activities and found that "*even a small amount of training can improve spatial reasoning*" (2013, p. 370). One of the interventions they reviewed is a 1 credit course that has been tried and tested over several years at Michigan Technological University (MTU) and emphasises mental transformation between 2 and 3 dimensional drawings in the style of isometric sketches, orthographic projections, coded plans and sectional drawings (Sorby, 2009). The course is attended by those who get a low score on a spatial test at orientation. It has been observed that completion of this course has led to improvements not only in spatial test scores but also to grades in a wide variety of subjects such as fundamentals of engineering and pre-calculus (Sorby & Veurink, 2010). In addition, higher retention rates have been measured among those who enrol in the course with women benefiting to a greater extent than men in this regard (Sorby, 2001).

This course has been successfully adopted by several other institutes in the US including a large number of engineering schools that participated in ENGAGE Engineering, a NSF funded project that started in 2009 ("ENGAGE Engineering," n.d.). The Department of Engineering Education at Ohio State University (OSU) was one of those who participated in ENGAGE and now includes the testing of both spatial and mathematical skills during freshman orientation. Therefore, engineering orientation at OSU is acting on findings from research such as those of Wai et al. (2009) and Sorby & Veurink (2010) and presents an opportunity to examine the relationship between spatial and math abilities and how they together predict performance in engineering education.

In this paper, we explore this relationship by addressing the following questions:

1. Do spatial skills predict success on a mathematics placement test routinely administered at the university for first-year engineering students?
2. Is there a gender gap in MPT scores and, if so, to what extent is this explained by the gender gap in spatial ability?
3. To what extent is one rewarded in assessment of course work for having good spatial skills?
4. Do spatial skills predict the ability to solve certain types of specific problems from the mathematics placement exam for first-year engineering students?

In this paper we present a statistical analysis of data from these tests and examine the extent to which these measures of spatial and mathematical ability correlate with each other, the types of math questions that reveal the highest correlation and how these correlations vary by gender.

Research Design

Data were collected for this study during the summer of 2016 using two instruments – The Purdue Spatial Visualization Test: Rotations (PSVT:R, Guay, 1976) and The Mathematics Placement Test (unpublished) developed at OSU. These tests are administered online as part of the OSU freshman orientation programme and described in more detail below. Approval for the study was obtained from the Institutional Review Board at OSU. At the time of taking the test, participants were provided with an IRB approved document containing information about this study and asked to consent to participate by allowing their responses to the tests to be included in the data set. A total of n = 1053 students provided informed consent and only their data is included in the study.

Instruments

The PSVT:R (Guay, 1976) consists of 30 multiple choice questions designed to measure 3-D mental rotation ability. One of the two practice questions provided at the start of the test is shown in Figure 2. This question involves the rotation of the object by 90° around the vertical axis. The participant must apply the same rotation to the second figure and select from one of the 5 options below a match of the rotated figure. There are 30 questions on the test with variation in the number of axes involved in the rotation and the type of figure the participant must mentally rotate. The test is timed so both speed and accuracy are assessed. Reliability measures for the PSVT:R are reported by (Yoon, 2011) with Cronbach's $\alpha = .81$

measured using data collected from a sample of 180 education major undergraduate students enrolled in mathematics courses.

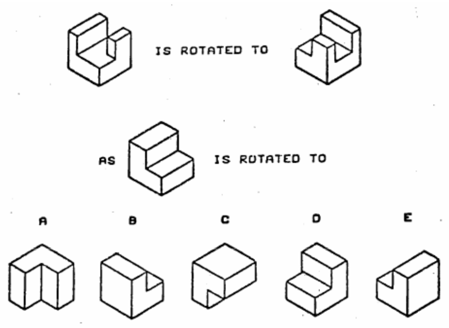


Figure 2. An example of a question on the PSVT:R (Guay, 1976).

The MPT was developed by the Department of Mathematics at OSU and is not publicly available. It consists of 25 questions, each of which has 5 variations that differ only in numbers used and, therefore, all 5 versions can be considered equivalent. Reliability was measured using Cronbach's α and found to be equal to .79 with none of the 25 items resulting in a higher value of α if removed. Therefore, the MPT is considered to have a high reliability. The topics included on the test are basic algebra, rational expressions, logarithm rules, exponent algebra, function notation, complex number arithmetic, inequalities, geometry and trigonometric rules and functions. All of the questions are multiple choice.

We also report two other measures of math ability, the Scholastic Aptitude Test (SAT) Math and the American College Testing (ACT) Math tests, and the ACT Science Reasoning test (SCIRE). Grade point average (GPA) scores gained by the sample in the Autumn semester, 2016, were also collected.

Method

Both the PSVT:R and the MPT were administered as part of the orientation process through the OSU learning management system in the summer immediately prior to enrolment. Participants took the tests online at a time and place of their choosing. Instructions were provided for each test. Times allowed for the PSVT:R and MPT were, respectively, 25 and 75 minutes. Sample problems were available for the MPT and the PSVT:R test contains two sample questions at the start.

Data Analysis

Based on the results of the PSVT:R, participants can be classified by spatial ability level as weak (PSVT:R ≤ 18), medium (19 to 21) or strong(≥ 22) (Veurink & Sorby, 2011). In this analysis we grouped both medium and strong visualizers as one group which we labelled 'strong' as this allowed us to compare two groups (weak and strong) without losing any of the data. Hence, for this study a weak visualizer is defined by PSVT:R ≤ 18 and strong as PSVT:R > 18.

The distributions of both the MPT and PSVT:R data were first examined and found to be skewed towards the upper end of their ranges, particularly so for the PSVT:R. This can be described as a ceiling effect with a large number of participants getting very high or maximum scores on the tests. This was more pronounced for the male compared to female cohort (skewness = -.930 for male vs. -.418 for female). However, since the sample size is large (n = 1053), the assumption of normality is presumed to hold in this case.

Statistical analysis consisted of grouping the sample by a criterion, e.g. gender, and comparing the scores of a variable, e.g. MPT, using an independent samples t-test to measure the significance of the difference and a Cohen's d effect size to measure the size of the difference.

Results

Descriptive statistics for the results from the MPT and PSVT:R tests are shown in Table 1 where they are presented for the entire sample and also grouped by gender. Included are the numbers of weak and strong visualizers in each group.

	Full Set		**Female**		**Male**	
	MPT	**PSVT_R**	**MPT**	**PSVT_R**	**MPT**	**PSVT_R**
n	1053	1053	272	272	781	781
Mean	18.15	23.58	17.29	21.55	18.45	24.28
Std. Deviation	4.234	4.506	4.198	4.855	4.208	4.155
PSVT:R ≤ 18		151 (17 %)		72 (26 %)		79 (10 %)
PSVT:R > 18		902 (83 %)		200 (74 %)		702 (90 %)

Table 1. Descriptive data for the full set and grouped by gender

These results show that 151 participants were eligible for voluntary enrolment on the spatial skills course offered in semester 1 at OSU. Performance on all of the tests are presented below and grouped first by gender (Table 2) and then by spatial ability level (Table 3).

Test	**Male**			**Female**			**t-test**	**Sig (2-tailed)**	**Cohen's d (Size)**
	n	**M**	**SD**	**n**	**M**	**SD**			
MPT	781	18.45	4.208	272	17.29	4.198	3.906	.000	0.28 (Medium)
PSVT:R	781	24.28	4.155	272	21.55	4.855	8.931	.000	0.61 (Large)
GPA	777	3.080	.738	271	3.074	.5864	.119	.905	0.01 (Small)
SAT Math	237	695.49	63.549	76	648.55	60.876	5.659	.000	0.76 (Large)
ACT Math	703	30.91	2.876	257	29.64	2.776	6.106	.000	0.45 (Medium)
ACT SCIRE	703	30.99	3.417	257	29.73	3.548	5.012	.000	0.37 (Medium)

Table 2. Differences in performance by gender on MPT, PSVT:R, GPA, SAT and ACT.

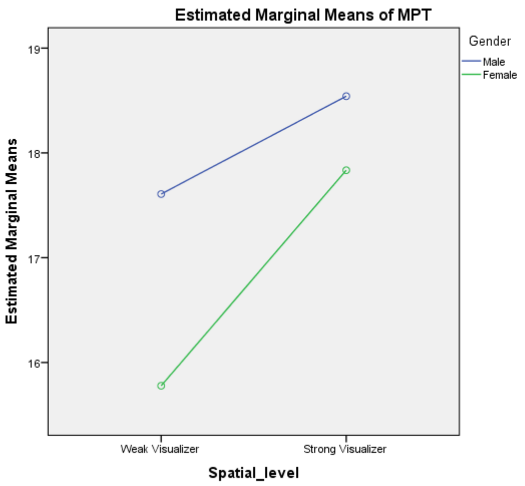
Test	n	Weak Visualizer M	Weak Visualizer SD	n	Strong Visualizer M	Strong Visualizer SD	t-test	Sig (2-tailed)	Cohen's d (Size)
MPT	151	16.74	4.29	902	18.38	4.18	-4.471	.000	0.39 (Medium)
PSVT:R	151	15.36	2.42	902	24.95	N/A			
GPA	148	2.84	0.67	900	3.12	0.70	-4.478	.000	0.41 (Medium)
SAT Math	39	664.10	62.69	274	686.93	66.04	-2.032	.043	0.36 (Medium)
ACT Math	134	28.92	2.58	826	30.84	2.87	-7.296	.000	0.71 (Large)
ACT SCIRE	134	29.22	3.45	826	30.89	3.45	-5.206	.000	0.49 (Medium)

Table 3. Differences in performance by spatial ability level on MPT, PSVT:R, GPA, SAT and ACT.

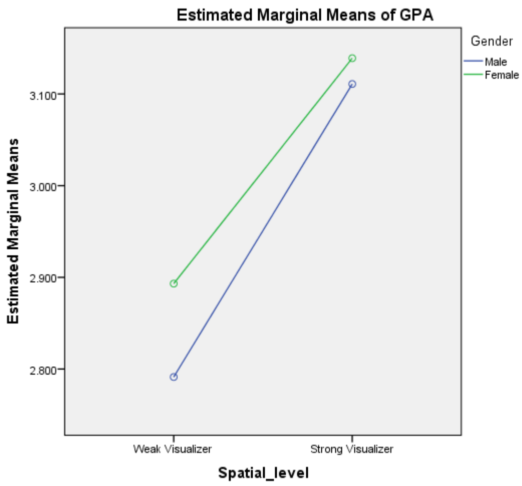
A two-way repeated measures ANOVA was conducted to measure the interaction of spatial ability level (weak or strong) with gender (male or female) on the different measures of academic performance collected from this sample (MPT, GPA, SAT, ACT, SCIRE) with results presented in Table 4 and Figure 3.

		Gender		Spatial ability		Gender x Spatial	
Variable	n	F	p	F	p	F	p
MPT	1053	11.180	.001	15.550	.000	2.194	.139
SAT Math	313	20.088	.000	1.649	.200	.426	.514
ACT Math	960	17.045	.000	38.505	.000	.503	.479
ACT SCIRE	960	11.208	.001	18.140	.000	.128	.720
GPA	1048	1.043	.307	19.717	.000	.335	.563

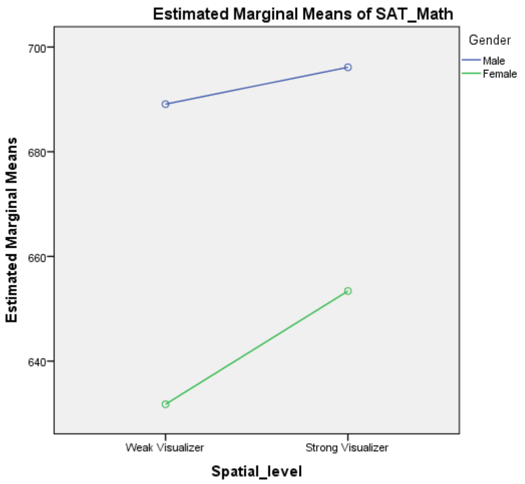
Table 4. Results of two-way repeated measures ANOVA to check for interaction between gender and spatial ability.



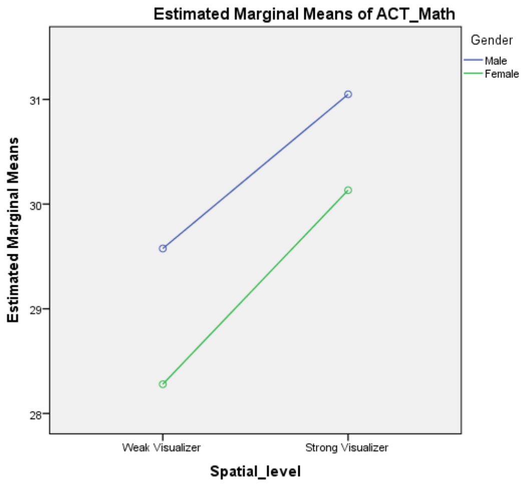
(a)



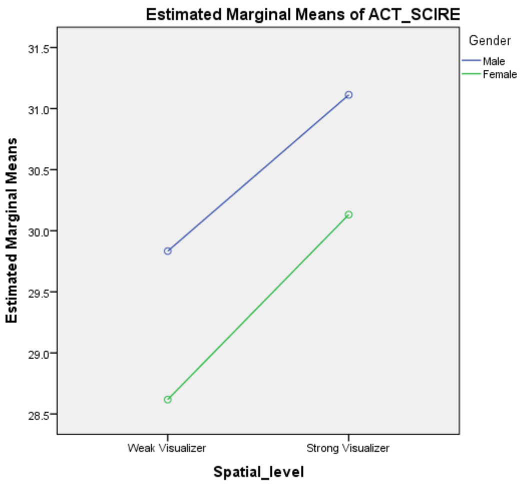
(b)



(c)



(d)



(e)

Figure 3. Graph of the interaction of gender and spatial ability level on (a) MPT, (b) GPA, (c) SAT Math, (d) ACT Math and (e) ACT SCIRE.

Finally, a correlation matrix is presented in Table 5 to show the extent to which each of the test measurements correlate with each other based on the full data set

	n	2	3	4	5	6
1. MPT	1053	.207**	.369**	.561**	.576**	.386**
2. PSVT_R	1053		.176**	.285**	.323**	.242**
3. GPA	1048			.264**	.413**	.302**
4. SAT_Math	313				.664**	.480**
5. ACT_Math	960					.506**
6. ACT_SCIRE	960					

* significant at p < .05, ** significant at p < .01.

Table 5. Correlation matrix for all tests.

Analysis of Results

The mean PSVT:R scores measured in this study for the entire sample, and separately for males and females are very similar to those reported for another group of US students (Sorby, Casey, Veurink, & Dulaney, 2013). The sample is also consistent with the internationally observed difference in spatial ability in favour of males (Lippa, Collaer, & Peters, 2010). In this case males scored significantly higher on the PSVT:R than females with the difference equal to 2.72 points and measured using Cohen's d to be a large effect size (d = .61, p < .01). Of the 272 females in the sample, 72, or 26 % of the females, were categorised as weak visualizers. In contrast, 79 of the 781, or 11 % of the males, were put in the same category. Hence, the weak visualizer cohort is 48 % female while the entire sample is 26 % female.

With regard to the first question, the extent to which spatial skills predict success on the MPT, the correlation between the two variables, although significant, is not very large (r(1050) = .207, p < .01). This result means that only a small amount of variation, 4.3 %, is shared between the PSVT:R and MPT scores. Hence, spatial skills, at least when measured by the PSVT:R, are not a strong predictor of success on the MPT. With regard to the other two math measures, ACT Math and the SAT Math, significant correlations were also measured but varied in magnitude. Ranked from smallest to largest correlation with the PSVT:R, the order is rPSVT:R-MPT(1050) = .207, rPSVT:R-SAT(311) = .285, rPSVT:R-ACT(958) = .323, all p < .01. A bigger gap between weak and strong visualizers is revealed by ACT Math compared with SAT Math and MPT. No significant difference in PSVT:R scores was measured between those who took the SAT Math (M=23.87, S.D.=4.457) and those who took the ACT Math (M=23.63, S.D.=4.445) tests, i.e. selection of either ACT or SAT was not influenced by spatial ability.

The second question relates to the issue of interaction between gender and spatial ability: is the relationship between spatial and math abilities different for males and females in this sample? The interaction between gender and spatial ability level was determined using a two-way repeated measures ANOVA, with each math measure separately entered as the dependent variable, which indicates there was no interaction between these variables on all three math measures (Table 4). The extent to which they interact is also shown in Figure 3 (a), (b) and (c) which illustrates males outperforming females in both spatial ability categories. In the case of the MPT, there is a noticeably larger difference between female weak and strong visualizers compared with male weak and strong visualizers. The correlations between the PSVT:R and MPT were measured separately for each gender to be rmale(779) = .166 and rfemale(270) = .228, both p < .01. Although there is no crossover on the two lines in Figure 3 (a) and the interaction between gender and spatial ability level was found to be not significant, there is a difference in magnitude of the correlations between the two measures for each gender with the correlation being higher for females. This leads to the female weak visualizer being ranked lowest in MPT scores (M = 15.78), followed by the male weak visualizer who is very close in MPT to the female strong visualizer (17.61 and 17.84, respectively) and highest is the male strong visualizer cohort with an average MPT of 18.54 (Table 6). Female weak visualizers are at the highest risk of being ranked poorly on the MPT.

Spatial ability	Male			Female			Δ MPT	t-test	Sig (2-tailed)	Cohen's d (Size)
	n	M	SD	n	M	SD				
Weak	79	17.61	4.307	72	15.78	4.088	1.83	2.671	.008	0.44 (Medium)
Strong	702	18.54	4.190	200	17.84	4.112	0.7	2.112	.035	0.17 (Small)

Table 6. **Comparison of MPT scores grouped by visualization category for male and female participants.**

The MPT is used to place students in an appropriate math course in first year engineering. Those scoring 8 points or lower are enrolled in algebra; scores of 9 to 12 lead to pre-calculus enrolment; and scores of 13 or higher results in the student being enrolled in calculus provided his/her high school grades are satisfactory and, if not, the student could be enrolled in pre-calculus. The data were next checked to determine how each group – male and female weak and strong visualizers – were distributed across these MPT categories. These results are shown in Table 7.

MPT_level	N	Mean	Std. Deviation	n weak male	n strong male	n weak female	n strong female
MPT <= 8	21	22.52	4.844	2 (3%)	13 (2%)	3 (4%)	3 (2%)
9 <= MPT <= 12	85	22.13	4.649	6 (8%)	45 (6%)	14 (19%)	20 (10%)
13 <= MPT <= 17	335	22.82	4.584	27 (34%)	211 (30%)	28 (39%)	69 (35%)
MPT >= 18	612	24.23	4.321	44 (56%)	433 (62%)	27 (38%)	108 (54%)
Total	1053	23.58	4.506	79 (100%)	702 (100%)	72 (100%)	200 (100%)

Table 7. **Distribution of sample by gender and spatial ability level across the four MPT levels**.

Although the correlation between the two measures, MPT and PSVT:R, is not large, the results in Table 7 show a difference in the number of students placed in each MPT category when grouped by gender and spatial level. Since the occupancies of the MPT ≤ 8 categories are very small they were grouped with the next category for analysis. Hence, 23 % of the female weak visualizers have a MPT ≤ 12 which is much higher than the equivalent percentages for the other groups. The distribution of male weak visualizers across the MPT categories is very similar to that of the female strong visualizers. As shown in Table 7, female weak visualizers have the greatest proportional representation in the algebra class (MPT ≤ 12) and this class is overrepresented by women who comprise 40 % of the class versus 26 % of the sample. The small correlation suggests the cognitive processes involved in answering questions on the MPT and the PSVT:R overlap to a small extent only but that, particularly for females, being classified as a weak visualizer predicts a higher risk of poor performance on the MPT. It has been suggested by others that the relationship between spatial and math skills may not be the same for males and females in that spatial ability plays a more important role in math performance for females (Halpern et al., 2007). Our data do not contradict this assertion.

Our third question relates to the relationship between performance in STEM education and spatial ability. Since it has been shown that measures of spatial ability and mathematics can explain more variance in academic achievement scores than mathematics alone we were interested in testing this hypothesis in our context and to also compare results by gender. GPA based on the first semester of first year was used as a measure of academic performance. During first semester, there is much variation in what courses are taken and contribute to GPA but the majority of first year engineering students are likely to take a course in math, general education and chemistry and one of four engineering courses. What contributes to GPA is, therefore, quite varied but it was the only measure of academic performance in semester 1 that was available. There was no significant difference in GPA by gender but a difference was found when the sample was grouped by spatial ability ($d = .41$, $p < .001$, Table 3). The relationship between spatial ability and GPA was very similar for both genders (Figure 3 (b)) and no interaction between spatial ability level and gender was observed for GPA (Table 4). The correlation between the PSVT:R and GPA was 3 % for male and 4 % for female. This is small – a change of 10 points in the PSVT:R score, which is large, is equivalent to a change of 0.3 in GPA.

While the correlation and effect size are small, a small change in GPA can have a big effect, particularly at borders such as GPA = 2.0 and GPA = 3.0 below which a student can be placed on academic probation, depending on the institute. Indeed, as shown in Figure 3, strong visualizers are on the right side of the GPA = 3.0 border. The sample was grouped into three GPA levels – GPA < 2.0, 2.0 ≤ GPA 3.0 and GPA ≥ 3.0 and the number of male and female weak and strong visualizers in each category was counted with the results are presented in Table 8.

GPA level	**n**	**Mean**	**Std. Deviation**	**n weak male**	**n strong male**	**n weak female**	**n strong female**
GPA < 2	83	23.04	4.964	8 (11%)	60 (9%)	5 (7%)	10 (5%)
2.0 <= GPA < 3	317	22.54	4.710	33 (43%)	188 (27%)	35 (49%)	61 (31%)
3.0 <= GPA < 4	648	24.19	4.235	35 (46%)	453 (65%)	32 (44%)	128 (64%)
Total	1048	23.60	4.504	76 (100%)	701 (100%)	72 (100%)	199 (100%)

Table 8. Distribution of sample by gender and spatial ability level across three GPA levels.

These results (Table 8) are slightly more favourable from the female perspective with proportionally greater representation of men in the lowest GPA category. The pattern in the highest GPA category is very similar for male and female weak visualizers and likewise for male and female strong visualizers. This is consistent with the comparison made in Table 2 above, i.e. when compared by gender there was no difference in mean GPA between male and female.

Our last question related to the nature of the questions on the MPT in terms of their relationship with spatial ability. We were interested in determining which questions explained the correlation between the two measures and why these questions drew more than others on spatial thinking. For each question in turn, we grouped the sample as having a correct or incorrect response to the question and then compared the spatial test scores of these two groups. We found a significant difference ($p < .01$) in the mean PSVT:R scores for 12 of the 25 questions but in many cases the difference was small. Only two questions had an effect size greater than 0.4 which equated to difference in mean PSVT:R scores between the correct and incorrect groups of 1.8 points. One of these questions was an algebra word problem related to mixing two components that required the translation of a problem statement into equations which are then solved for two unknowns. The other required the comprehension of an inverse relationship between two variables. Thirteen of the 25 problems revealed no difference in spatial ability, i.e. the mean PSVT:R score of the correct and incorrect groups were equal. These were very much plug and chug type operations in which an equation was provided and procedure had to be followed. These questions tested core competencies in mathematics at a procedural rather than a problem solving level. One simply had to adopt the schema provided in the question and follow the rules.

Discussion

The small correlation we found between MPT and PSVT:R suggests an inconsequential relationship between the two variables, yet we found some noticeable and meaningful differences in the occupancy of MPT categories by male and female weak and strong visualizers. Female weak visualizers were overrepresented in the algebra class and the average MPT score of this group places it inside the 13 to 17 range that does not guarantee enrolment in the calculus course. The policy at OSU is to check high school grades for MPT scores in this range. Female students tend to perform well on high school grades (Halpern et al., 2007) and are, therefore, likely to compensate with these grades and be placed in the calculus course with a MPT score in this range.

When compared to the men, women begin the first semester at OSU with a large deficit in spatial skills, SAT Math and ACT Math and a slightly lower MPT score but finish the semester with equal grades. Assuming GPA draws on verbal, mathematical and visuospatial abilities, at a cognitive level, and on several emotional and affective aspects such as motivation, work ethic and so on, there are at least two possible explanations for this observation: (i) the male and female suite of cognitive abilities remain unchanged relative to each other but females compensate for other reasons or (ii) females improve their math and verbal abilities to reach parity with men and all other things are equal. Since our data set is limited to the cognitive ability profile collected at the start of the semester we do not have sufficient information to answer this question. However, since GPA starts accumulating from day 1 of the semester, which started with a difference in ability profile by gender, we suggest the answer lies closer to the former explanation in which females compensate in other ways for a shortfall in math and spatial abilities.

Compared to the PSVT:R scores, the correlations between GPA and the three math measures (MPT, SAT-Math, ACT-Math) are larger. Ranked by increasing order they are $r_{SAT\text{-}Math}$ =

.264, r_{MPT} = .369, and $r_{ACT\text{-}Math}$ = .413 (all $p < .01$), with the corresponding r^2 values being .07, .14 and .17. Using linear regression with GPA as the dependent variable, one of the three math measures was entered as the first dependent variable with PSVT:R as the second in order to measure the additional change in r^2. This was repeated separately for each math measure. PSVT:R was found to improve the regression for the MPT only and by only 1 %. In other words, the PSVT:R scores had a negligible effect as a second variable to mathematics in predicting GPA. Since correlations between spatial ability and STEM assessments have been found where the assessments are non-routine (Duffy, Sorby, Nozaki, & Bowe, 2016), a plausible explanation is that semester 1 assessment tasks are closely related to course material.

All MPT questions came with multiple answers from which the participant made one selection. Having multiple answers to select from can facilitate a trial and error approach which must help some students in getting the correct answer. In other words, removing or not providing a selection of answers will lead to greater demands on cognitive abilities, greater variation in the MPT scores and, possibly, a higher correlation between spatial and MPT. Where spatial ability has been found to correlate with mathematical skills it is typically when participants are presented with non-routine tasks. The highest correlations found in a meta-analysis by Friedman (1992) of 136 studies reporting mathematical-spatial correlations was for mathematical reasoning tasks. Casey, Nuttall, & Pezaris (2001) found that 8th grade girls outperformed boys in tasks such as multiplying fractions but were outperformed by the boys when asked to perform tasks such as estimating the height of a tree based on the known height of a person standing beside it. We have found a small correlation between the PSVT:R and the MPT, a test which is of the multiple choice format and whose majority of questions are designed to assess core competency and routine mathematical tasks. One finding from this study is, therefore, that spatial ability is of little need when mathematics is assessed in this way.

Conclusions

The mean PSVT:R scores we measured in this sample are typical for US engineering students. While correlation with the MPT was small, PSVT:R scores did have some power in predicting the relative positioning of male and female weak and strong visualizers using PSVT:R = 18 as the weak/strong cut off point. There was no significant interaction between spatial ability level and gender on any of the measures, including MPT. GPA in semester 1 was best predicted by the ACT-Math score with spatial ability adding little to the correlation. Many of the questions on the MPT did not place demands on spatial thinking; those that did were less procedural and more open-ended. While there was little correlation between GPA and PSVT:R, mean GPA of strong visualizers was greater than 3.0 while for weak visualizers it was less than 3.0. It is also possible the curriculum post first year will contain more non-routine assessment tasks that place more varied demands on cognitive abilities. If so, larger correlations and patterns such as those observed by Wai et al. (2009) may be observed.

References

Casey, M. B., Nuttall, R. L., & Pezaris, E. (2001). Spatial-mechanical reasoning skills versus mathematics self-confidence as mediators of gender differences on mathematics subtests

using cross-national gender-based items. *Journal for Research in Mathematics Education*, 28–57.
Duffy, G., Sorby, S., Nozaki, S., & Bowe, B. (2016). Exploring the role of spatial cognition in problem solving. In *Frontiers in Education Conference (FIE), 2016 IEEE* (pp. 1–4). IEEE. Retrieved from http://ieeexplore.ieee.org/abstract/document/7757593/
ENGAGE Engineering. (n.d.). Retrieved February 6, 2017, from http://www.engageengineering.org/
Friedman, L. (1992). *A Meta-Analysis of Correlations of Spatial and Mathematical Tasks*. Retrieved from eric.
Guay, R. (1976). *Purdue spatial visualization test*. Purdue University.
Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest*, *8*(1), 1–51.
Lippa, R. A., Collaer, M. L., & Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Archives of Sexual Behavior*, *39*(4), 990–997.
Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, *93*(3), 604.
Sorby, S. A. (2001). Improving the Spatial Skills of Engineering Students: Impact on Graphics Performance and Retention. *Engineering Design Graphics Journal*, *65*(3), 31–36.
Sorby, S. A. (2009). Educational Research in Developing 3-D Spatial Skills for Engineering Students. *International Journal of Science Education*, *31*(3), 459–480. https://doi.org/10.1080/09500690802595839
Sorby, S. A., Casey, B., Veurink, N., & Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learning & Individual Differences*, *26*, 20–29. https://doi.org/10.1016/j.lindif.2013.03.010
Sorby, S. A., & Veurink, N. (2010). Long-term Results from Spatial Skills Intervention among First-Year Engineering Students. Presented at the Proceedings of the 65th Midyear Meeting of the Engineering Design Graphics Division of ASEE.
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.
Veurink, N., & Sorby, S. A. (2011). Raising the Bar? Longitudinal Study to Determine Which Students Would Benefit Most From Spatial Training. Presented at the Proceedings of the ASEE 2011 Annual Conference and Exposition, Vancouver, BC, Canada.
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.
Yoon, S. Y. (2011). *Psychometric Properties of the Revised Purdue Spatial Visualization Tests: Visualization of Rotations (The Revised PSVT-R).* ERIC. Retrieved from http://eric.ed.gov/?id=ED534824
Yorke, M., & Longden, B. (2008). The first-year experience of higher education in the UK. *Higher Education Academy. Available at Http://www. Heacademy. Ac. uk/assets/York/documents/resources/publications/FYEFinalReport. Pdf [Accesses on 2 September 2010] THE INTERNATIONAL JOURNAL OF INTERDISCIPLINARY SOCIAL SCIENCES*.