Practice Report: Six Studies of Spatial Skills Training in Introductory Computer Science

By Jack Parkinson, *University of Glasgow*, Ryan Bockmon, *University of Nebraska-Lincoln*, Quintin Cutts *University of Glasgow*, Michael Liut, *University of Toronto Mississauga*, Andrew Petersen, *University of Toronto Mississauga* and Sheryl Sorby, *University of Cincinnati*

We have been training spatial skills for Computing Science students over several years with positive results, both in terms of the students' spatial skills and their CS outcomes. The delivery and structure of the training has been modified over time and carried out at several institutions, resulting in variations across each intervention. This article describes six distinct case studies of training deliveries, highlighting the main challenges faced and some important takeaways. Our goal is to provide useful guidance based on our varied experience for any practitioner considering the adoption of spatial skills training for their students.

INTRODUCTION

The relationship between spatial skills and computer science is not a newly discovered connection, with research in this area appearing occasionally since the 1950s [17], but interest in the area is growing. Several researchers have shown that actively training spatial skills can have an impact on computing outcomes [2,3,11]. Since spatial skills training is lightweight and easy to deliver compared with many computing exercises, and since it may open hitherto unrecognised learning pathways, spatial skills training may well be of interest to many researchers and practitioners. This article details six implementations of spatial skills training courses at several institutions, with various changes across each instance. We share our experiences, highlight the challenges faced and recommend some methods of training spatial skills for computing students. The aim is to make others considering spatial skills training for their cohort aware of the many possible degrees of freedom we explored to help them in determining the most appropriate program for their own context.

WHAT ARE SPATIAL SKILLS?

'Spatial skills' is an umbrella term that captures several distinct but independent skills all related to extracting information from "flat" representations, like paper or a screen, and constructing robust internal visualizations of structures or objects. One might assume that they are all about navigating and visualizing real-world space, like when you need to parallel park or re-arrange the furniture in a room—while these tasks are probably not completely divorced from spatial skills, the skills we are talking about in this article are more abstract and cognitive. Some factors of spatial skills include mental relations (understanding how objects in a space relate to each other), perceptual speed (identifying a known pattern from an unobscured environment) and closure flexibility (identifying a known pattern from an obscured environment).

They are inherently internal, and usually are quantified by psychometric tasks. Mental rotation is one of the tasks that comes under the factor of *spatial visualization*, one of the umbrella spokes that make up spatial skills. Such a task might involve a participant to determine a sequence of rotations that would transform one view of a 3D object to match another view, then apply the same sequence to another object. There are several cognitive tasks involved: identifying the 3D structure of objects from a 2D representation, formulating and holding the 3D structure in your head, and then making changes to the 3D structure in-situ. These are different skills, but all contribute towards spatial skills generally.

'Spatial skills' is an umbrella term that captures several distinct but independent skills all related to extracting information from "flat" representations, like paper or a screen, and constructing robust internal visualizations of structures or objects.

SUMMARY OF THE RELATIONSHIP BETWEEN SPATIAL SKILLS AND COMPUTER SCIENCE

Two theories for the relationship between spatial skills and computing have been published. Both Parkinson and Cutts [10] and Marguliеux [7] propose that spatial skills are not strictly necessary for succeeding in computing—or STEM—but that they do help. Both theories present the idea that spatial skills develop abstract skills that are valuable in tackling general problems and tasks associated with STEM fields. For example, Parkinson and Cutts suggest that spatial relations (a sub-skill of spatial ability) aids in moving between levels of Schulte's Block Model [13], and Margulieux suggests that repeated spatial skills practice develops encoding strategies for non-verbal information.

Both theories are grounded in research but yet are not fully tested or explored, so our understanding of spatial skills and their influence on STEM skills is not complete. The first time that spatial skills training was run by the authors, not much was known about the relationship between spatial skills and computing science. We were aware of evidence supporting the following:

- Spatial skills correlate with success in multiple areas of computing [4,5,10]
- Spatial skills correlate with attainment in academic computing [10]
- Spatial skills training appeared to impact outcomes in the computing electives of engineering students [15]
- Spatial skills training appeared to improve the CS assessment outcomes of students in a summer school [3]
- Spatial skills are a more powerful intervening factor than access to computers when examining computing achievement in the context of socio-economic status [9]

Since the interventions began, we now know more about the relationship, with research indicating:

- Spatial skills training positively impacts outcomes in distinct deliveries of introductory computing programs [2,11]
- Spatial skills are associated with low level expression evaluation [12]
- Spatial skills can, particularly when combined with other factors, predict student success in a CS1 program [1]

SPATIAL SKILLS TRAINING COURSES: CASE STUDIES

Although the deployment and delivery of spatial skills training courses differs per implementation, the course content itself is fundamentally the same (with one caveat: Case 6 did not include exercises on paper or cube folding, since the intention of Case 6 was to replicate Cooper et al.'s work in the area as closely as possible, and Cooper did not originally have time to include paper folding in the instruction [3]). It is comprised of Sorby's training course: *Introduction to Spatial Visualization* [14,16]. This course has been shown to be effective for engineering students across several years of use so we were confident that, regardless of the effect on computing outcomes, it should at least be effective in training spatial skills.

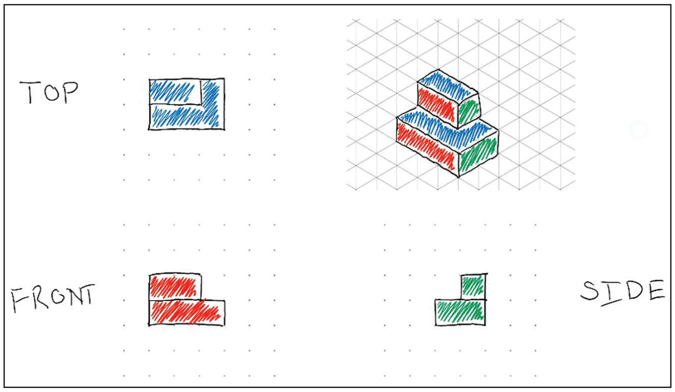

Figure 1: A representation of an isometric view of an object (top right) and the three orthographic views of the object (labelled Top, Front and Side). NB: color is used here for clarity, and usually would not be visible to students except in examples and explanatory resources. A typical spatial skills training exercise may require a student to draw out one or more orthographic views from an isometric view, choose the correct corresponding orthographic view(s) from a multiple-choice selection, draw an isometric view from orthographic views, identify whether a given set of orthographic views is correct (i.e., are lined up correctly and form a geometric shape) and so on. These are a sample of some of the types of exercises Sorby uses to develop spatial skills.

The course has nine chapters consisting of exercises involving orthographic and isometric projection, 3D rotation about axes, reflection, symmetry, flat patterns, and combinations of solids. The chapters are split into a total of 37 exercises and roughly 10-20 questions per exercise. The exercises are drilling exercises, with sequences of questions in similar styles designed to enforce concepts over consecutive attempts. Many consist of multiple-choice questions requiring participants to circle an image or letter option, a substantial number are drawing exercises with a guiding grid provided, and some require text- or symbol-based answers to be written. Through repeated drilling on the exercises, students practice visualization in different contexts and thereby develop their spatial skills. Several examples of questions can be found throughout the article.

Another constant that remained unchanged throughout was the test of spatial skills, the Revised PSVT:R [18]. This is a test of mental rotation which requires a participant to: view two orientations of the same object; determine the set of rotations required to achieve the second orientation given the first as a starting point; apply the set of rotations to another object; and identify the resulting orientation from a selection of five possibilities. The test is timed at 20 minutes and consists of 30 items in this style, arranged in increasing difficulty (that is, later items require assimilation of more rotations about more axes). Where applicable, we also consistently utilised Sorby's score breakpoints: any score below 19 is a fail (training is likely necessary), a score between 19 and 21 is a marginal pass (training is likely to be beneficial) and any score above 21 is a solid pass (training is not likely to be required) [14].

Finally, all studies consider only students in introductory computing programs, specifically in their first year of study. This is for two main reasons: firstly, Sorby's course has been

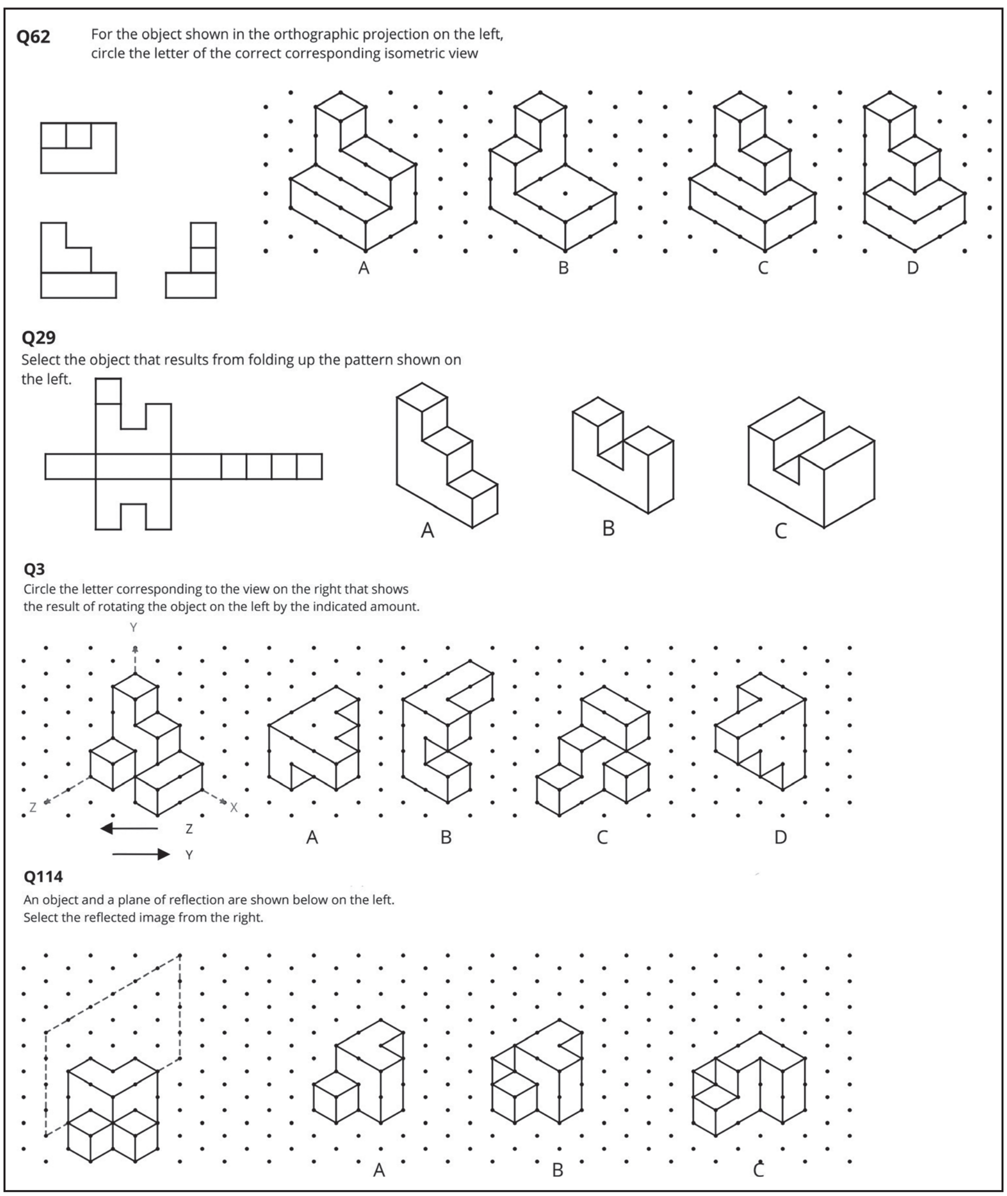

Figure 2: A broad selection of multiple-choice questions from the exercises. Answers are C, A, B, and B respectively.

designed for engineering students in their first year of study, so in each case study it was deemed sensible to focus on this cohort. Secondly, Margulieux's theory for spatial strategy states that spatial skills are of most value to novices beginning their higher education journey, so these students were likely the ones who needed the training the most.

There were variations made across several different factors in each case study. These were:

- **Participants:** who took the course and how were they selected.
- **Timing:** when the course was delivered during the institution's first semester.
- **Delivery method:** this varied between using pencil-and-paper exercises that were marked by an instructor, exercises through a virtual learning environment (VLE) that were automatically marked, exercises through a bespoke platform that were automatically marked, and combinations of these platforms.
- **Recognition of participation:** sometimes completion of the course was formally recognized; other times participation was just voluntary. Note that since some of the cases were not credit bearing, *mandatory* is not necessarily the best term to use. However, for the purposes of this paper, we use *voluntary* to indicate, "come along and do this course if you want to," and *mandatory* to indicate, "the school expects you to take part in this course and it will be recognised on your transcript." Any deviation from these broad definitions of voluntary and mandatory will be made clear in the text.
- **Scheduling:** this indicates how the work was timetabled each week, with options being formally timetabled (as in, appearing in the students' official university timetables) or *informally* timetabled (there was a set time to attend sessions but these were communicated via emails and forum posts as typical extra sessions are, therefore not appearing in students' official university timetables) and *fully timetabled* (that is, every hour of work is accounted for), *semi-timetabled* (some hours are accounted for in sessions but students are expected to fit some work into their schedule) or *not timetabled* (there are no set sessions and students are expected to fit all the work into their schedule).
- **Support:** this represents the communication channels by which students were supported in their learning.

Margulieux's theory for spatial strategy states that spatial skills are of most value to novices beginning their higher education journey, so these students were likely the ones who needed the training the most.

Some of these case studies are based on training courses that ran during major research projects into the relationship between spatial skills and computing science, and therefore have been summarized and published at various venues. Typ-

Table 1: Variations in factors of interest in each case study.

	Participants	Timing	Delivery	Recognition	Scheduling	Support
Case 1	CS0 students, based on a PSVT:R taken under exam conditions	Weeks 6-10	Paper based and VLE (Moodle)	None; voluntary	Informally semi-timetabled	In-person, email
Case 2	CS0 and CS1 students, based on a PSVT:R taken online	Weeks 4-8	Paper based and VLE (Moodle)	None; voluntary	Informally fully timetabled	In-person, email
Case 3	CS0 and CS1 students, based on a PSVT:R taken online	Weeks 4-8	Paper based	0-credit PASS/FAIL on transcript; mandatory	Formally timetabled	In-person
Case 4	CS0 and CS1 students, based on a PSVT:R taken online	Weeks 3-7	Online SST platform	0-credit PASS/FAIL on transcript; mandatory	Not timetabled	Course forum, tutorial videos, MS Teams messaging
Case 5	All CS1 students	Weeks 1-8	Online SST platform	Contributing towards 5% of final CS1 grade	Not timetabled	Course forum, tutorial videos, email
Case 6	CS1 students, invited to participate	Weeks 3-16	Paper based and VLE	Extra course credit or $10 gift card	Not timetabled	Video lectures, example videos and online examples and practice

ically, though, they lack focus and detail on exactly how the training was run, just enough to make it clear to the audience how the course was delivered. The aim of this article is to revisit these projects—and explore some that haven't been published—to break down exactly how training was conducted and discuss the staff and student experiences.

For this reason, some of the cases don't include substantial detail on the effect of the training with respect to computing outcomes—these outcomes are already published, so it isn't necessary to reiterate them in detail here. It also lets us focus on the main outcome of this article—the experiences we have had with training and recommendations on how to conduct this training with computing students.

Some of these case studies [presented] are based on training courses that ran during major research projects into the relationship between spatial skills and computing science, and therefore have been summarized and published at various venues.

CASE 1: VOLUNTARY, VLE & PAPER

The focus of Case 1 was a cohort of CS0 students (that is, students taking introductory computing science with little-to-no prior programming experience) who were tested immediately after one of their computing assessments in the middle of their first semester. The test was taken in exam conditions during a session when almost 100% attendance was recorded, so we were able to capture the spatial skills of almost the whole cohort. Of 101 students, 35 failed and 19 marginally passed. Based on these scores, spatial skills training was recommended to the 54 students scoring a fail or a marginal pass. They were invited in class and via an email from the program convenor to attend scheduled sessions, with a brief overview on the existing work connecting spatial skills and computing science as motivation to participate.

The course was split into drawing exercises and multiple choice or short answer exercises that were hosted on the institution's VLE (Moodle). The VLE exercises were automatically marked, to be done online in the students' own time with a recommendation that they spent one hour on them per week. Each week the instructor met the students for an hourly drawing session to briefly introduce the work, supervise the session, and check in with students to see that they were comfortable with the drawing and the online content. The instructor's email address was also available for questions on work done outside of scheduled hours.

Of the 54 invited students, 18 attended at least one in-person session and only four attended all the sessions. At the end of the course, five students took a spatial skills post-test. One student's score decreased by one point, and the other four students showed substantial increases in their scores (an improvement of 12 points in one student). Two pieces of computing assessment—one immediately prior to testing and one at the end of the semester—were used to determine any significant gains in computing aptitude after training. The students were ranked as a class according to their assessment scores in each test, and their class rankings were compared. Those that took training—though a small group—jumped up in the class rankings compared to those who did not, showing a statistically significant gain. Notably, no students who took training went down in class ranking between the two assessments.

CASE 2: VOLUNTARY, VLE & PAPER

In Case 2 the participants and testing were slightly different. Rather than halfway through the semester, training was started sooner, with the expectation that the sooner students could develop their spatial skills, the more they would be able to apply them in their computing modules. The test was issued to both the CS0 and CS1 cohorts at the authors' institution as part of their introductory lab with the intention of offering training to both cohorts.

Two hundred and fifty students took the initial test. Of these, 78 (31%) failed and 40 (16%) marginally passed. These 118 students were invited to participate in training in a similar fashion to Case 1, with an email invitation indicating prior research, but no extrinsic motivation or incentive offered. Delivery was also similar, except that rather than expecting the online component to be completed in their own time, students were allocated two hours per week in the computer labs to complete both the drawing and the VLE exercises. Again, an instructor was available to support during the scheduled sessions.

Twenty-eight students attended at least one training session and six students attended every session. Five students took the PSVT:R post-test, showing a mean gain of six points. Since the training was started earlier, there were no previous assessments that could be used to calculate gains across the computing program.

CASE 3: MANDATORY, PAPER ONLY

In Case 3, the course was made mandatory for students across CS0 and CS1 scoring below the 19-point threshold of the PSVT:R, meaning that all students were required to take the spatial skills test as part of their introductory lab to determine whether they would be allocated to training. This resulted in 225 students across CS0 and CS1 taking the spatial skills test with 64 (28%) requiring training. A further 41 (18%) marginally passing students invited to take part but not required—these students are discounted for the remainder of this Case description since only four students opted to take part, so the number of students enrolled for the purposes of this article was 64.

The course began in week 4 after a mid-term assessment. It was entirely paper-based, with all the questions compiled into weekly workbooks. The course was also formally timetabled, with two 1-hour sessions scheduled per week appearing on the students' official timetables with all their other modules, located in a flat floored study space with a flexible, open seating arrangement. Each session was supervised by an instructor who provided on-the-spot feedback and some light touch instruction as students worked. The instructor also marked students' workbooks between sessions and handed them back at the start of each week.

Fifty-three students completed the training course and took a spatial skills re-test at the end; the remaining 11 students did not attend at all, so the course had effectively no drop off but did have no-shows (though it is possible that these students switched out of computing science and did not inform the spatial skills convenor). The completing students saw a mean gain of about six points in the PSVT:R, in line with other case studies, and in general also showed gains in their computing measures above the marginally passing students. More detailed results for this study can be found in work by Parkinson and Cutts [11].

CASE 4: MANDATORY, ONLINE SST PLATFORM

A fourth case arose because of the COVID-19 pandemic that restricted in-person access to campus spaces and made a paper-based course virtually impossible. Instead, a bespoke online spatial skills training platform (henceforth referred to as the online SST platform, to be distinct from an online VLE platform) was developed to host both the drawing and the more simple-answer questions. The platform had automatic marking, some generic feedback mechanisms, and tutorial videos, meaning that most of the content could be completed entirely through the platform. A Microsoft Teams instance was also created with an FAQ and a question forum, allowing interaction between students and Q&A with the instructor.

The initial test was delivered as in Case 3: it was a compulsory part of the introductory lab. Again, the course was made mandatory for students who scored below 19, however it was not formally timetabled since there was an expectation that students would fit the exercises in around their work schedule. The course was started as early as possible, in week 3, after the initial add/drop period had ended. The exercises were split into roughly-equal weekly blocks across five weeks, to give an indication of what students should be working on each week, but students weren't prohibited from skipping ahead and doing extra work.

Three hundred and fourteen students took the initial test; 84 (27%) scored below the threshold and were enrolled in the course. Of these 84, 54 completed the course entirely (completing every exercise and taking the post-test), 18 students began the work but did not complete, and 12 students did not take part at all. Once again, gains of about six points in the PSVT:R were observed as a result of training and in general the students with high spatial skills scored highly in their computing assessment (regardless of the route taken: some were initially high scorers, and some improved their scores via training). Additionally, students who completed training—regardless of their original scores—were among the highest scoring students in their CS assessment and almost indiscernible from students who begin the program with high spatial skills. There may be some bias in these results, though: students who did not take part in the training may just be students who are not engaging with the program at all, explaining their low results.

CASE 5: ENTIRE COHORT, ONLINE SST PLATFORM

The fifth case was conducted in partnership with the authors at a different institution. The same online SST platform was used, however students from the entire CS1 cohort—substantially larger than the previous institution's cohorts—were all expected to do the training, not just the ones with initially poor spatial skills. The decision was taken to make all students participate because spatial skills are transferrable and would likely be of value to all students, and there were internal issues relating to requiring some of the cohort to do extra credited work when some weren't expected to.

The completion of the spatial skills training contributed to 5% of the students' final marks in the CS1 module (although it was made clear that all other assessment would not include spatial skills tasks themselves). The course was initially planned to span six weeks with two hours work per week, but this was extended to eight weeks after students said that the spatial skills tasks took too long to complete in two hours per week.

Table 2: Computing assessment scores (in %) for both CS0 and CS1 categorised by students' original spatial skills score and how much spatial skills training they completed. Those originally passing were not invited to take part in training. Note that the highest assessment scores are consistently achieved by those who completed training, which are also like the scores of the students who originally did not need training.

	CS0: did not take training	CS0: partially completed training	CS0: completed training	CS1: did not take training	CS1: partially completed training	CS1: completed training
Originally FAIL	49.5 (n=10)	63.2 (n=12)	73.9 (n=29)	46.0 (n=2)	48.3 (n=6)	59.2 (n=25)
Originally MARG	71.4 (n=12)	69.0 (n=2)	75.5 (n=5)	52.0 (n=37)	59.0 (n=1)	59.0 (n=2)
Originally PASS	75.0 (n=50)	N/A	N/A	61.7 (n=121)	N/A	N/A

The students had a course Q&A forum that was used to issue questions and concerns about the spatial skills exercises, as well as an email address built into the online interface. Students were also asked to provide feedback on the platform each week through the tool they used to do their weekly work.

In this study, students were given a pre- and post-test in spatial skills. Of 668 students enrolled in the CS1 module, 467 students (70%) completed the spatial skills training course and 434 (65%) took both the pre- and post-test. A mean gain of 2.5 points in the PSVT:R was observed after training. However, when we consider only the students who a) completed all the training and b) initially scored below 19 points on their PSVT:R pre-test, the number of students is reduced to 121 and the gain increases to a mean of 6.4, that aligns with other case studies. More detailed results for this study can be found in work by Ly et al. [6].

The most interesting case is Case 5, in which the whole cohort was required to take training.

CASE 6: VOLUNTARY, VLE AND PAPER

The sixth case was conducted across three other institutions over a period of two years. The first year was focused on collecting baseline data, with the PSVT:R and the SCS1R [8] (a reduced set of the language independent programming test for CS1 students, the SCS1) issued on a voluntary basis to all CS1 students to measure spatial skills and programming abilities respectively. The students were encouraged to participate with different incentives at each institution: extra credit to students' final grade, extra credit for a test or a $10 gift card. The tests were conducted in the first two weeks of the 16-week semester with 274 respondents from the three institutions (from a combined pool of approximately 800 CS1 students).

In the second year of the study a spatial skills intervention was run with the results measured against the baseline data from the previous year. Students were again tested during the first two weeks of the semester with both the PSVT:R and the SCS1R. The intervention used eight of Sorby's nine chapters and was taught over 12 weeks starting in the third week of the semester. Students were tested again at the end of the semester with both the PSVT:R and the SCS1R to see if there were any gains in their abilities over the semester. The material for each module consisted of several online video lectures, an optional online practice and hands-on 3–4-page worksheets to be completed and handed in over the module's 1–2-week period. Worksheets had to be printed, then scanned and turned in online by being sent to a designated email address to be marked. Marking was carried out by an instructor at each institution using the same marking key.

All students who were enrolled in a CS1 module at each of the universities were asked to participate in the second-year study. All participation was voluntary with the same incentives offered as to the students in the first year (a form of credit or a gift card). A total of 71 students participated in the study. Results showed that running the spatial skills intervention increased students' spatial skills and programming abilities compared to students who did not participate in the study. Although the gains were small (only 1.2 points in the PSVT:R on average and 2.0 points in for students scoring 18 and below in the PSVT:R) there was a strong correlation between the students' post-PSVT:R and post-SCS1R. More detailed results for this study can be found in work by Bockmon et al. [2].

DISCUSSION

In this section we break down the factors that were varied in each case study and examine their implications. For each factor, we make our recommendation on how to overcome the issues. Note, however, that these recommendations are based only on the case studies described above, so it's possible that they may only be specific to the institutions examined. More studies and more analysis are required to determine whether these recommendations are universal or specific only to the institutions included in the article.

EFFECT ON COMPUTING OUTCOMES

In all instances, spatial skills training was observed to yield at least marginally positive outcomes for computing students. In Case 1 and Case 2, the participation was so low that we should only really consider these case studies as learning experiences for the instructors to better understand how students interact with the training and what should be adjusted in future iterations. Case 4 has not featured in any publication, though some light-touch analysis of the results that would seem to indicate that the training was beneficial are included; though once again, the best value this study provides is data on the first deployment of the online SST platform. The remaining cases—3, 5 and 6—have featured in peer reviewed publications each of which provides its own analysis of the benefits of the training in a CS context.

PARTICIPANTS AND PARTICIPATION

Each case study had some issues with encouraging student participation. Many students do not appear to be interested when the course is not mandatory, with the best non-mandatory results observed in Case 6 (roughly 9% participation, with incentive provided). Making the course mandatory increases participation, but also increases pushback: students often wonder why they need to do these tasks that have the appearance of being unrelated to their computing education.

The most interesting case is Case 5, in which the whole cohort was required to take training. The pushback appears to be higher than in other case studies, although this may appear inflated for two main reasons: 1) the volume of students was high so the volume of complaints/concerns will of course also be high, and 2) students were given the opportunity to report on their experiences weekly to indicate how they felt about it, something not strictly encouraged in any other case.

Regardless, the reason for the stronger pushback may be that we expected students who already had good spatial skills to do the work. As indicated by the difference in gains for this case study, it is not likely that the course was of any real benefit to them, but they were still expected to spend several hours working on it. This probably bred resentment, that may have "spread" among the cohort through adjacent conversations and student spaces, perhaps souring the experience for students who otherwise would have been neutral. This is mostly speculation, but it is certainly worth considering whether we need to put students through extra work regardless.

Our recommendation is to only include participants in training who really need it. This may raise some ethical concerns about giving some students more work than others, but our justification is that it is comparatively low effort when considering a typical introductory CS module and the potential benefits evidenced by some recent studies appear to outstrip the time-cost to students.

TIMING

We believe that starting the course earlier will be of more benefit to the students in the long term: they will have more time to allow these skills to develop and will be able to apply them across more of the course. However, starting early might make it difficult to report on the effects of a training course with respect to computing outcomes: the effectiveness of any training is hard to gauge without any meaningful pre-tests of computing, that are challenging to administer early in the semester. Without measures of effectiveness and evidence that the training is beneficial, it is hard to justify the intervention, both to students and to institutional leads.

We expect that there are two options: one could trust that the existing research is evidence enough to indicate that training spatial skills is valuable for CS students and opt to begin the intervention as soon as possible to maximally benefit students, or one could attempt to support research efforts by replicating previous work and ensure that pre-tests are used to measure CS aptitude, perhaps in the form of a mid-term test or some other computing assessment.

Regarding the length of the course, we have not observed any adverse effects from condensing the course into five weeks. In most case studies a gain in PSVT:R score of about 6 points was observed, that is in line with Sorby's findings when delivering the course over a longer period of time, and with Case 5's slightly longer delivery (when considering the subsection of students who most likely needed—and completed—the training). Case 6 is an outlier with the longest time being spent on the course and the lowest gains observed.

DELIVERY PLATFORM

Delivery is the most complex and nuanced of the factors discussed, with variations ranging from ease of use, feedback mechanisms, and student attitudes.

The instructor of the course for Cases 1 and 2 observed that using a VLE is not an effective way of delivering the course content. It will functionally serve for multiple choice questions, but for more complex interactions the VLE becomes clunky and unintuitive. See Figure 2 for an example of a simple, intuitive interaction on paper that becomes more complex and convoluted in a VLE. Moreover, see Figure 3 for an example of a drawing exercise, that cannot be reasonably recreated in this institution's VLE at all, meaning that any attempt to use a VLE will likely require the course to be split across paper and online delivery.

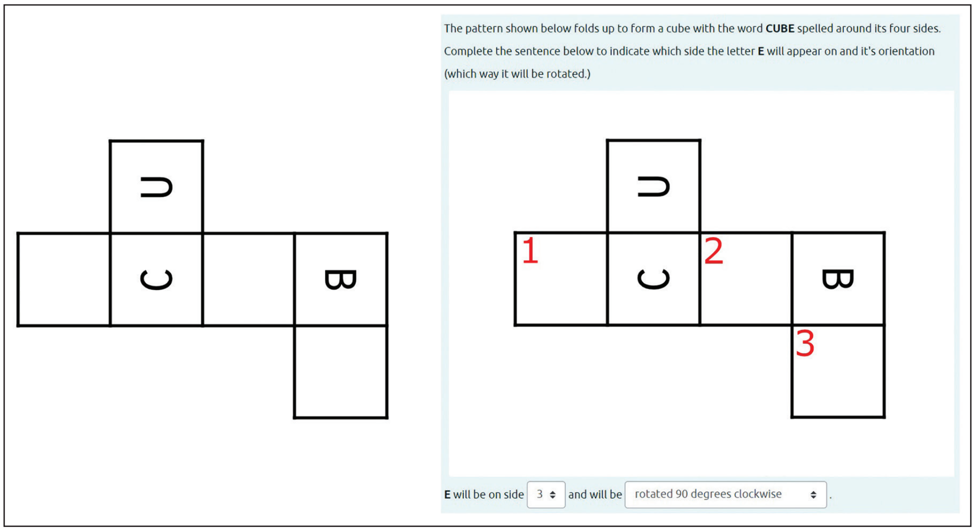

Figure 3: A reproduction of a question similar to one from Sorby's workbook (left) where the participant must add the missing letter with the correct orientation such that when the flat pattern is folded into a cube, it spells out "CUBE" around the outside faces; and the implementation of this question (right) in Moodle, the VLE used, with two input fields and labels to identify the possible inputs. Turning the page and writing an "E" in the correct orientation is trivial compared to selecting multiple answers from dropdown menus to correspond to numbered cells.

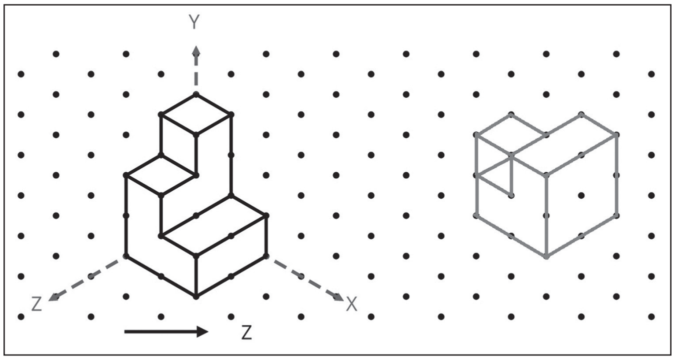

Figure 4: A reproduction of a question similar to one from Sorby's workbook requiring the student to apply a rotation (given as an arrow code) to an object and draw the resulting orientation on the isometric grid provided. Arrow codes are one method of denoting spatial operations that are explained during the course: an arrow pointing right indicates a rotation of 90 degrees anti-clockwise, and Z denotes the axis to rotate the shape by. The answer can be seen in grey on the right.

Many of these interaction issues were overcome by developing the online SST interface. The complexity of the interaction in Figure 2 was overcome by allowing the user to click in a box and have the missing letter appear. Clicking in another box would move the missing letter there and clicking in a box that already contained a letter would rotate it by 90 degrees so that the student could adjust the orientation.

The online SST platform and the VLE allow for automatic marking, which is both a blessing and a curse. On the surface, it allows students to immediately check their work and see where they are going wrong so that they can quickly adjust their process without having to wait for an instructor; or worse, continuing to complete questions incorrectly without realising. On the other hand, it allows less motivated students to guess and click random answers until they are correct without being penalised, reducing the effectiveness of the work. Students in Case 5 were asked to answer Likert questions after training was complete indicating how much they had guessed on the course, and 18% indicated that they almost always or often guessed answers, with a further 36% indicating that they sometimes did. There are ways to mitigate this: for example, only allowing students one free "check" per question or applying a forced delay between checks, but neither of these are likely to be pleasant for the students. With some careful logging, one could also record the times between checks and feedback to the student when it looks like students are guessing based on the patterns of frequency. However, each of these mitigation strategies require an arbitrary breakpoint somewhere, so it is impossible to develop a strategy that will not capture false positives or negatives in every case.

Students in both Cases 4 and 5 also indicated that there were issues with the platform. Issues reported were usually relating to errors in the system or difficulties in interacting with the interface. The drawing interface, while still considered by the authors to be the best drawing interface for this kind of exercise we have encountered in a web app, can sometimes be unintuitive or can demonstrate strange functionality when it comes to marking due to its complex nature. Checking that lines are drawn in the correct place may seem straightforward, but in practice, since a user can begin a drawing from any point on the grid, the calculation required to check the lines must be dynamic and flexible by many degrees. Usually, the calculation was successful in marking incorrect lines in red and correct ones in green, but sometimes if the placement of the object was slightly "off," it could report many of the lines as wrong when generally the structure was correct, giving the students the wrong idea of how correct their work was. An example of this is shown in Figure 4.

An in-person paper-based system appears to be the best solution, especially if the instructor is there to give feedback on students' work as they go. However, the instructor for Case 3 struggled to mark and provide meaningful feedback for 50 students per week and was not always able to complete the marking in time, so while feedback and guidance were

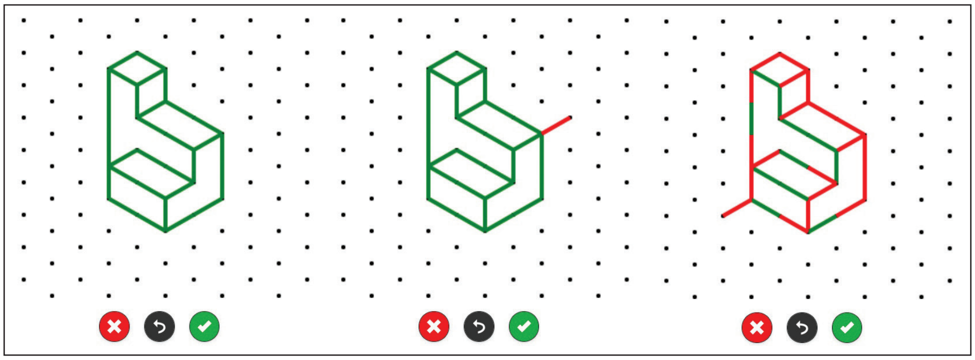

Figure 5: An example of a case where a single misplaced line can have a marked effect on the reported "correctness" of the drawing. From left to right, the first image shows the correct drawing; the second image shows the correct drawing with one additional line, highlighted in red (indicating that it is an incorrect line); the third image again shows the correct image with a single misplaced line, however the reported correctness is very different due to the way the lines are checked. Since there is no specific starting location, there is a very large pool of potentially correct answers, so we cannot simply check if a line is present in "the right place."

generally rich, nuanced, and personalized, they were not always possible to effectively deliver. The instructors for Case 6 had no problems with marking their proportions of the 71 workbooks, which they each had two weeks to mark due to the longer course delivery.

Based on general feedback from all the cases, Case 3 was clearly outstanding as the most satisfying and enjoyable student experience. Students were far more likely to indicate that they were engaged and interested in the content. The instructor verified this, as he witnessed more peer discussion and general curiosity about the exercises in the in-person drawing classes when students sat around large tables facing each other. There is some additional nuance about being in that environment, versus looking at screens or even drawing in front of screens, concerning the promotion of a shared learning experience. Students will pass around the eraser, glance over to see how their peers are doing, pick up and show their workbooks to each other, even provide assistance by making small pencil marks on each other's work to try and demonstrate motion or visibility of objects. All these small interactions promote collaboration and geniality, that students said made the course enjoyable.

Interactions like these were less common in Cases 1 and 2 and were virtually non-existent in Case 4. In fact, more than one Case 4 student said that they felt isolated and alone, the antithesis of Case 3. This could be partially alleviated, along with the scheduling issue (see below), by having scheduled online sessions in which students all come online to work together, perhaps even with a shared voice or video channel. This would at least promote discussion and sharing of progress although it's unlikely that the flexibility of passing notebooks or drawing marks on each other's work can be replicated in an online setting.

Based on all these findings, we found paper-based, in-person sessions to be the best-received by students and the best way to provide meaningful feedback, though this was challenging for the instructor and for a large cohort might be impossible to scale. It is possible that an online SST platform could be modified to mitigate some of the issues faced in Cases 4 and 5 by carefully tracking students' interactions and observing when they are not engaging with the material properly, as well as supplementing their learning experience with community-based sessions, however it is unlikely to reach the same level of nuanced, personalised delivery as the former. Although no explicit feedback was collected to gauge the student experience, Case 6 presents a possibility of conducting paper-based work while distanced, which may be of use to some practitioners too.

Based on all these findings, we found paper-based, in-person sessions to be the best-received by students and the best way to provide meaningful feedback, though this was challenging for the instructor and for a large cohort might be impossible to scale.

RECOGNITION

The case studies where the course was mandatory saw much higher completion rates, even in cases where they did not actually bear any credit or make a difference to the students' progression. By formalising the content and making it clear that it was an expectation of the institution that students take part, many of them saw this as a signal that the course should be completed.

SCHEDULING

The way cases were scheduled had a substantial impact on students' attitudes towards them. Cases 1 and 2 had a schedule, but it was not officially included in their timetables; uptake was low, but when both a drawing session and an in-person VLE session were scheduled, students were much more likely to properly attempt both. In Case 1 where only the drawing sessions were scheduled and the VLE content was to be completed in their own time, the VLE content was all but ignored by most students, even those who attended more than one drawing session. By adding another hour to the schedule in Case 2 specifically to work on the VLE content in the computer labs, this content was attempted by many more students, even if ultimately very few went on to complete the course.

A schedule that is formally timetabled, however, was clearly the most effective method. Attendance at each session was very high, and even when students could not attend, they generally made efforts to prove that they had done the work.

It was thought that by not providing a schedule for the online content in Cases 4 and 5, students would fit it in whenever was most convenient for them, since it was still compulsory. This was met with substantial resistance, with many students claiming that they simply did not have enough time to fit in two hours of spatial skills training per week. This was particularly interesting for Case 4, since in their academic year they had less formal contact time than the previous cohort described in Case 3—when it is not scheduled it appears that the students push back against what is perceived as "extra" work. Most of these students were placated by the instructor walking them through their timetable and demonstrating that there is space for some additional work, but some students denied that it was possible to fit into their timetable and simply did not complete the course.

If for no other reason than to give the perception of an important piece of work, it looks like officially timetabling sessions is a valuable step to take in formalising the intervention, regardless of the format in which it is delivered. It is the instructor's opinion that if Case 4 had been officially timetabled, as many of the students' other online CS modules were, there would have been much less resistance and push-back against the extra work.

MOTIVATION

An observation was also made relating to intrinsic student motivation and helping them to understand why such an intervention is important. Ideally, students would be participating because they believe that the course would help them, not just to check a box to make their institution happy or to earn some credit, but it can be challenging to instill this intrinsic motivation. The consistently used approach was to provide evidence that training is likely to be beneficial based on prior research into the relationship. Unfortunately, despite this being made clear at the start of each intervention, many students respond in feedback across each case study that they don't really understand why they are doing the intervention or how it is related to CS. The most successful case in this regard was Case 3, but the instructor believes there is some bias here because while he was present and overhearing conversations between engaged students about spatial skills, he would often chip in to cite relevant prior research and contribute to their discussions. This continual reference to the evidence, introduced organically as part of an ongoing discussion, likely had a lasting impact on the students' perceptions of the importance of the training.

Increasing evidence is accumulating that supports the connection between spatial skills and computing success, so we encourage readers to consider developing the spatial skills of their students, and we offer up the experiences and recommendations in this article as a starting point.

PSVT:R DELIVERY

Although this has not been explored in detail, we have some concerns about the reliability of the PSVT:R in different delivery contexts. It is a psychometric test that should be completed individually and require the students' full attention, but this is not something we can guarantee in any case study beyond Case 1, when both tests were conducted in exam conditions. Evidence to suggest that we should be concerned was gathered in Case 5, when some students actually took two post-test PSVT:Rs, one immediately after training and one two weeks later as the semester wrapped up. We expected that the scores of these two tests would be very similar for each participant because it was unlikely that their spatial skills would have changed over two weeks without any direct intervention, but a non-negligible standard deviation was observed.

We could not say exactly why this occurred. As stated, Case 1 had tests conducted in exam conditions; Cases 2 and 3 had tests conducted in a lab when students were expected to be working individually, but the lab may have been busy and full of distractions, and the post-tests were conducted in the students' own time; in Cases 4, 5 and 6, all tests were conducted in the students' own time whenever and wherever they saw fit. This means we have a wide variation of conditions under which the tests were taken, a variation that could threaten the validity of the results in some of the case studies. Note, however, that in each instance (whether taken in a lecture theatre or through an online interface) the time limit of the PSVT:R has always been 20 minutes with examples and instructions provided prior to the timer starting.

Our recommendation would be to conduct the PSVT:R in exam conditions on paper where possible, to reduce threats to validity such as students being distracted, students sharing or discussing answers, images not loading correctly on an online platform and for the more nuanced reason that students are more likely to take it seriously.

LIMITATIONS

Concerns have been raised about the suitability of the training course for low- or no-vision participants, as well as students with non-verbal learning disabilities. In all the case studies examined here, the only related concern was color-blind students, who indicated that the red-green marking scheme was difficult to follow (this was addressed by providing detailed textual feedback for each question, and a version of the online application is being developed using color-blind friendly colors). Without any experience of students with visual or non-verbal disabilities, we are unable to address these issues directly. This is an important concern that warrants much closer examination.

RECOMMENDATIONS

To summarize our discussion of the six cases above, we make the following broad recommendations.

- The sooner the intervention can begin, the better. This gives students more time to develop the skills and apply them in their studies, based on existing research.
- Only provide training to those who need it—best practice so far is to use Sorby's PSVT:R breakdown scores.
- In-person, paper-based exercises in flat-floored workspaces where students can sit together appears to be the most enjoyable and stimulating format for students (however, take note of the difficulty in providing feedback and marking work, especially for large cohorts). Online platforms could be used provided students are given an opportunity to develop community bonds and potential misuse of the platform is accounted for, however a typical VLE probably won't have enough complexity of interaction to accurately recreate the tasks.
- Uptake is substantially greater when the course is formally recognised and delivered as a compulsory course by the institution, even when the recognition of student

participation does not have any impact on progression or grades.
- Formal scheduling, with the intervention sessions appearing in students' official timetables, is more likely to be better attended and less likely to receive push-back from students.
- Intrinsic motivation should be developed by presenting prior research that shows the connection between spatial skills and success in computing science, and ideally this should be part of an ongoing conversation throughout the course.
- The PSVT:R test is best conducted on paper in exam conditions.

It is worth noting once again that all the training courses are essentially—as far as we can tell—reasonably effective in terms of training spatial skills, as measured by gains in the PSVT:R. Therefore, if any of the proposed case studies were used as a basis for an intervention it is likely that it would be beneficial to the students, but our recommendations are based on the many other factors involved.

CONCLUSION

There is no single method of developing spatial skills that, as of now, can be determined to be objectively the best, specifically in the domain of CS education. However, we have learned a lot about what works well and what doesn't over the years of delivering a spatial skills training course; enough that we feel that our recommendations are of use to any researchers or practitioners considering getting involved with spatial skills instruction. Increasing evidence is accumulating that supports the connection between spatial skills and computing success, so we encourage readers to consider developing the spatial skills of their students, and we offer up the experiences and recommendations in this article as a starting point. ❖

References

1. Bockmon, R., Cooper, S., Gratch, J., Zhang, J., and Dorodchi, M. Can Students' Spatial Skills Predict Their Programming Abilities? in *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, (ACM, New York, NY, 2020): 446–451. doi: 10.1145/3341525.3387380.
2. Bockmon, R., Cooper, S., Koperski, W., Gratch, J., Sorby, S., and Dorodchi, M. A CS1 Spatial Skills Intervention and the Impact on Introductory Programming Abilities, in *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, (ACM, New York, NY, 2020): 766–772. doi: 10.1145/3328778.3366829.
3. Cooper, S., Wang, K., Israni, M., and Sorby, S. Spatial Skills Training in Introductory Computing, in *Proceedings of the Eleventh Annual International Conference on International Computing Education Research*, (ACM, New York, NY, 2015): 13–20. doi: 10.1145/2787622.2787728.
4. Jones, S. and Burnett, G. Spatial Skills and Navigation of Source Code, in *Proceedings of the 12th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, (ACM, New York, NY, 2007): 231–235. doi: 10.1145/1268784.1268852.
5. Jones, S. and Burnett, G. Spatial Ability and Learning To Program, *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 4, 1 (2008), 47–61. doi: 10.17011/ht/urn.200804151352.
6. Ly, A., Parkinson, J., Cutts, Q., Liut, M., and Petersen, A. Spatial Skills and Demographic Factors in CS1, presented at the Koli Calling 2021, Joensuu, Finland, Nov. 2021. doi: 10.1145/3488042.3488049.
7. Margulieux, L. E. Spatial Encoding Strategy Theory: The Relationship between Spatial Skill and STEM Achievement, in *Proceedings of the 2019 ACM Conference on International Computing Education Research*, (ACM, New York, NY, 2019): 81–90. doi: 10.1145/3291279.3339414.
8. Parker, M. C., Guzdial, M., and Engleman, S. Replication, Validation, and Use of a Language Independent CS1 Knowledge Assessment, in *Proceedings of the 2016 ACM Conference on International Computing Education Research*, (ACM, New York, NY, 2016): 93–101. doi: 10.1145/2960310.2960316.
9. Parker, M. C., Solomon, A., Pritchett, B., Illingworth, D. A., Marguilieux, L. E., and Guzdial, M. Socioeconomic Status and Computer Science Achievement: Spatial Ability as a Mediating Variable in a Novel Model of Understanding, in *Proceedings of the 2018 ACM Conference on International Computing Education Research*, Espoo Finland, Aug. 2018, 97–105. doi: 10.1145/3230977.3230987.
10. Parkinson, J. and Cutts, Q. Investigating the Relationship Between Spatial Skills and Computer Science, in *Proceedings of the 2018 ACM Conference on International Computing Education Research*, (ACM, New York, NY, 2018): 106–114. doi: 10.1145/3230977.3230990.
11. Parkinson, J. and Cutts, Q. The Effect of a Spatial Skills Training Course in Introductory Computing, in *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, (ACM, New York, NY, 2020): 439–445. doi: 10.1145/3341525.3387413.
12. Parkinson, J., Cutts, Q., and Draper, S. Relating Spatial Skills and Expression Evaluation, in *United Kingdom & Ireland Computing Education Research Conference*. (ACM, New York, NY, 2020): 17–23. doi: 10.1145/3416465.3416473.
13. Schulte, C. Block Model: An Educational Model of Program Comprehension as a Tool for a Scholarly Approach to Teaching, in *Proceedings of the Fourth International Workshop on Computing Education Research*. (ACM, New York, NY, 2008): 149–160. doi: 10.1145/1404520.1404535.
14. Sorby, S. and Baartmans, B. A Course for the Development of 3-D Spatial Visualization Skills., *Engineering Design Graphics Journal*, 60, 1, (1996): 13–20.
15. Sorby, S., Veurink, N., and Streiner, S. Does spatial skills instruction improve STEM outcomes? The answer is 'yes', *Learning and Individual Differences*, 67 (2018): 209–222. doi: 10.1016/j.lindif.2018.09.001.
16. Sorby, S., Wysocki, A. F., and Baartmans, B. *Introduction to 3D Spatial Visualization: An Active Approach*, 1st ed. Delmar Cengage Learning, 2003.
17. Super, D. E. and Bachrach, P. B. *Scientific Careers and Vocational Development Theory: A Review, a Critique and some Recommendations*. (Teachers College, Columbia University, New York, NY, 1957).
18. Yoon, S. Y. Psychometric properties of the Revised Purdue Spatial Visualization Tests: Visualization of Rotations (the Revised PSVT:R), Purdue University Dissertation, 2011.

Jack Parkinson
University of Glasgow
Glasgow, UK
jack.parkinson@glasgow.ac.uk

Michael Liut
University of Toronto Mississauga
Toronto, Canada
michael.liut@utoronto.ca

Ryan Bockmon
University of Nebraska-Lincoln
Lincoln, USA
rbockmon@cse.unl.edu

Andrew Petersen
University of Toronto Mississauga
Toronto, Canada
andrew.petersen@utoronto.ca

Quintin Cutts
University of Glasgow
Glasgow, UK
quintin.cutts@glasgow.ac.uk

Sheryl Sorby
University of Cincinnati
Cincinnati, USA
sorbysa@ucmail.uc.edu

DOI: 10.1145/3494574