

## **Which methodologies can be used to identify engineering students' spatial skills?<sup>1</sup>**

### **Quelles méthodologies peut-on utiliser pour repérer les habiletés spatiales d'étudiants ingénieurs?**

### **Quais metodologias podem ser usadas para identificar as habilidades espaciais de estudantes de engenharia?**

**Sophie Charles**

ID ORCID : 0000-0002-4499-5842

*Laboratoire Euler, ISAE-Supméca*

KEY WORDS: human activity analysis, solving strategies, spatial skills, spatial tests

*Research on spatial ability is interested in identifying the cognitive processes which characterise it. It has been related to success in science, technology, engineering, and mathematics. Completed as part of the French research project EXAPP\_3D e-FRAN, our study has focused on measuring engineering freshmen's spatial and 3-D modelling skills. Pen-and-paper tests are currently the most popular method used to measure spatial skills. Our research has shown that performance on these tests is not indicative of the skills targeted by the tests alone: Our first subjects reported a variety of solving strategies. The sensitivity of spatial tests to bypass strategies, in the sense that those used do not correspond to the skills targeted by the test, presents the challenge of defining an experimental protocol that characterises both the performance measured in these tasks, as well as the skills they engage. This article presents a mixed methodology which addresses this twofold issue.*

---

1 The French version was published in issue 47(3) 2024: <https://doi.org/10.7202/1119633ar>



**MOTS CLÉS :** analyse de l'activité, habiletés spatiales, stratégies de résolution, tests spatiaux

*La recherche sur les habiletés spatiales s'intéresse à qualifier les processus cognitifs qui la caractérisent. Dans le cadre du projet e-FRAN EXAPP\_3D, notre recherche s'est portée sur la mesure des habiletés spatiales et des compétences de modélisation volumique d'étudiants ingénieurs primo-arrivants. Pour ce faire, nous avons investigué la mesure des habiletés spatiales la plus fréquemment observée dans les études contemporaines, c'est-à-dire les tests papier-crayon. Notre recherche a mis en évidence que la performance relevée dans ces tests n'est pas révélatrice des seules compétences visées : nos premiers entretiens ont révélé une pluralité de stratégies (cognitives, comportementales) mobilisées. Cette sensibilité des tests spatiaux aux stratégies de contournement présente le défi de définir un protocole expérimental qui permette de caractériser aussi bien la performance mesurée dans ces tâches que les compétences qui y sont mises en œuvre. Cet article présente la méthodologie mixte que nous avons conçue pour répondre à ce double enjeu.*

**PALAVRAS-CHAVE:** análise da atividade, estratégias de resolução, habilidades espaciais, testes espaciais

*A investigação sobre as habilidades espaciais procura qualificar os processos cognitivos que as caracterizam. No âmbito do projeto e-FRAN EXAPP\_3D, a nossa investigação centrou-se na medição das habilidades espaciais e das competências de modelação volumétrica de estudantes de engenharia recém-chegados. Para isso, investigámos o método de avaliação das habilidades espaciais mais frequentemente utilizado nos estudos contemporâneos, ou seja, os testes de papel e lápis. A nossa investigação revelou que o desempenho observado nesses testes não reflete exclusivamente as competências visadas: as primeiras entrevistas mostraram uma pluralidade de estratégias (cognitivas e comportamentais) mobilizadas pelos estudantes. Esta sensibilidade dos testes espaciais às estratégias de contorno coloca o desafio de definir um protocolo experimental que permita caracterizar tanto o desempenho medido nessas tarefas como as competências efetivamente mobilizadas. Este artigo apresenta a metodologia mista que concebemos para responder a este duplo desafio.*

Authors' note: Correspondence regarding to this article may be addressed to [sophie.charles@isae-supmeca.fr](mailto:sophie.charles@isae-supmeca.fr)

## Introduction

This article aims to present a methodology for investigating spatial abilities more heuristic than simple performance assessment, by supplementing spatial performance measures with an investigation of engaged cognitive processes. Observation of multiple and alternative strategies used in spatial tests (Hegarty, 2018; Pilardeau, 2008; Workman et al., 1999) raises the question of what they actually measure. According to some researchers (Barratt, 1953; French, 1965; Guay, 1980; Guay et al., 1978; Lohman, 1979, 1996; Thurstone, 1938), the use of strategies not targeted by certain spatial tests is reason to question their validity. If a test allows the use of alternative strategies, is it possible to be sure of what it really measures?

Since Galton's work on imagery (1883), research on spatial abilities has focused on describing the cognitive processes they engage (Eliot, 1983; Tartre, 1990). Traditionally, these abilities have been measured using performance tests and paper-and-pencil tests (Eliot, 1983). Paper tests are more widely used, as they are considered easier to use (Lohman et al., 1987), than performance tests, which require the manipulation of materials (Lohman et al., 1987) or movement performance (American Psychological Association, 2018c). Use of computerised tests for assessing these skills has increased in recent years (Branoff & Dobelis, 2012; Kelly Jr, 2013; Onyancha et al., 2009; Tsutsumi et al., 1999), as they allow for automated management of answers and response time measurement (Branoff, 2000).

Spatial skills have been specifically linked to success in science, technology, engineering and mathematics (STEM), both for success in studying these disciplines and career orientation in these specialties (Wai et al., 2009). Studies specific to engineering training programmes have established the predictive nature of spatial tests for performance in volumetric modelling tasks (Branoff & Dobelis, 2012; Steinhauer, 2012), i.e. computer simulation of an object in three dimensions (Lieu & Sorby, 2009).

As part of the French e-FRAN investment programme, which focuses on the positive and negative effects of the use of digital technology in teaching and learning practices (Ministry of National Education, Higher Education and Research, Ministry of Economy, Industry and Digital Affairs, and General Commission for Investment, 2016), our research (Charles, 2023) explored the measurement of spatial skills and volumetric (3-D) modelling abilities of first-year engineering students. The primary objective was to use tests to study the link between performance in 3-D modelling and spatial performance, with performance understood as a result obtained by collecting responses (American Psychological Association, 2018b). This work examined the prerequisites for engineering students who learn 3-D modelling in design studies, training necessary for professional practice (Ault & John, 2010; Geronimi et al., 2005; Hamon, 2009). The first phase of our work consisted in identifying, obtaining, and experimenting tests designed to measure various spatial factors. Our initial experiments revealed that the participants used several problem-solving strategies to solve the tasks presented in the tests. In light of this confirmation of previous findings by Hegarty (2018) and Pilardeau (2008), we designed a methodology for characterising spatial performance and problem-solving strategies. For the data collection, we combined quantitative and qualitative methods, putting first-year engineering students through a battery of five spatial tests to assess their performance, then asking participants to complete open and closed questions on the strategies they had used to observe the cognitive processes at work.

## Conceptual framework

### *Spatial ability*

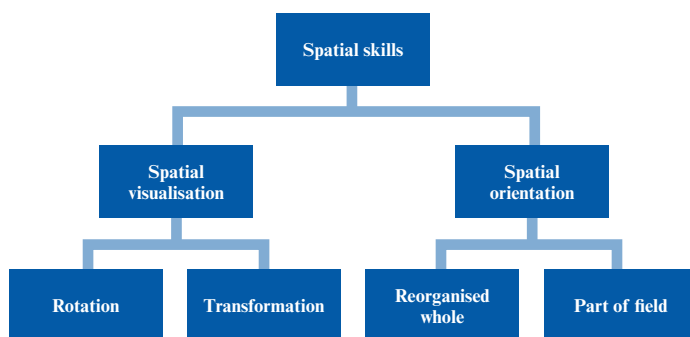
According to Lohman (1996), spatial ability refers to the ability to “generate, retain, retrieve, and transform well-structured visual images” (p. 98). No consensus has been reached on the nature or number of skills it requires (Uttal et al., 2013), but several authors (Hegarty, 2010; Linn & Petersen, 1985; Maier, 1996; Tartre, 1984; Uttal et al., 2013) have proposed classifications. Drawing on eleven models in the existing literature, we observed the recurrence of two main factors (Charles, 2023): Spatial visualisation and spatial orientation. The first factor, spatial visualisation, relates to the mental manipulation of two- and three-dimensional objects (Linn & Petersen, 1985; McGee, 1979; Tartre, 1984). The second factor, spatial

orientation, refers to a person's ability to determine the orientation of an object from a given point of view (McGee, 1979; Tartre, 1984). Thus, in the case of spatial visualisation, the observer mentally manipulates an object, whereas in the case of spatial orientation, it is a matter of "comprehending or making sense out of a visual representation" (Tartre, 1984, p. 28).

Tartre's classification (1984), illustrated in Figure 1, uses this distinction to categorise tasks designed to measure spatial skills according to the cognitive process they are intended to mobilise (Tartre, 1990). The author reiterates the distinction made by Kersh & Cook (1979, cited in Tartre, 1984) between the skill of manipulating an entire object, known as mental rotation, and that of manipulating part of an object, known as mental transformation. For example, a mental rotation task consists in mentally rotating a shape, whereas imagining folding a pattern to make a cube is a mental transformation.

Similarly, Tartre (1984) separates the ability to "understand a visual representation" or "some change that has taken place between two representations" (p. 6) from the ability to understand the relationship between an object and the whole from which it originates. For example, understanding the two possible interpretations of Necker's cube (1832) requires a change of perspective, whereas placing a piece in a jigsaw puzzle requires understanding the relationship between the piece and the pattern of which it is a part.

Figure 1  
*Tartre's spatial skills classification scheme (1984)*



*Note.* Simplified representation of Tartre's spatial skills classification (1984, p. 27).

### *Measuring spatial ability*

Results of psychometric tests were explored to identify the spatial factors used in solving spatial problems and hence define spatial ability (Eliot, 1983). This section presents five tests devised to measure the four spatial factors derived from Tartre's classification (1984). The process for this selection will be described in the methodology.

The first three tests focus on spatial visualisation. The Mental Rotation Test (MRT) (Vandenberg & Kuse, 1978) and the Revised Purdue Spatial Visualisation Test: Rotations (R PSVT:R) (Yoon, 2011) aim to measure mental rotation. The Mental Cutting Test (MCT) (College Entrance Examination Board, 1939) intends to assess mental transformation. The last two tests focus on spatial orientation: the Purdue Spatial Visualisation Test: Visualisation of Views (PSVT:V) (Guay, 1976a) targets the ability to understand an object regardless of how it is represented, and the Closure Flexibility Test (Concealed Figures) Form A (CFT) (Thurstone & Jeffrey, 1965) the ability to isolate an element embedded in a complex pattern.

The MRT presents 20 questions in which the respondent must identify the two orientations of an object that are identical to the one presented as a stimulus from among four possibilities. Respondents must provide two correct answers per question to earn one point (Vandenberg & Kuse, 1978), with a maximum score of 20 points. This test is organised in two parts, separated by a break. The duration of the test must be adapted according to the experience of the respondents and the test battery used (A. R. Kuse, personal communication, 25 June 2018).

The R PSVT:R presents 30 questions for which the participant must identify the correct answer from five possibilities. To do so, they are asked to observe the sequence of rotations applied to an example to use it on the stimulus presented in the question. Each correct answer is worth one point, the highest score is 30 points. The response time is unlimited in order to measure the respondents' actual level of spatial ability (Yoon & Mann, 2017).

The MCT presents 25 questions in which the respondent must choose from five alternatives to identify the section resulting from the cut shown on the stimulus. As each correct answer is worth one point, the highest score is 25 points. The response time is limited to 20 minutes (Eliot & Macfarlane Smith, 1983).

The PSVT:V presents 30 questions in which an object is depicted. The respondent must imagine this object floating in the middle of a glass cube. A black dot, positioned on one of the corners of the glass cube, indicates the viewpoint to be adopted to imagine the representation of the stimulus from this angle. The respondent must choose between five possibilities the view of the stimulus determined by the viewpoint indicated on the glass cube. Each correct answer is worth one point and the highest score is 30 points. The response time is limited to 20 minutes (Eliot & Macfarlane Smith, 1983).

The CFT presents 49 questions in which the respondent must indicate whether the stimulus is embedded in four complex patterns. For some questions, there are several correct answers. The score is calculated by adding the number of correct answers and subtracting the number of incorrect answers. The maximum score is 196 points. The time limit is ten minutes, and the instructions specify that respondents are not expected to answer all the questions within the time limit (Thurstone & Jeffrey, 1956).

### *Characteristics of spatial tests and affordance*

The format of the tests described above have common attributes: they are multiple-choice questionnaires (MCQs) with a time limit (with the exception of the R PSVT:R). MCQs have been described as promoting comparison and selection among possible answers (Leclercq, 2006; Lohman, 1979) and not encouraging the production of an independent answer (Bloom & Broder, 1950; Hopkins, 1998; Leclercq, 2006). Tests without time limits are considered to allow for the implementation of a greater number of strategies (Schwartz, 1963), whereas tests with very tight time limits, whose instructions specify that answers must be correct despite this constraint, would modify the participants' response style (Hopkins, 1998). For example, Cooke-Simpson and Voyer (2007) observed that female respondents, faced with the dual constraint of responding quickly and correctly, preferred to not respond rather than respond randomly.

Furthermore, the shapes used in the questions have specific characteristics. In the PSVT:V, the R PSVT:R and the MCT, the shapes are isometric representations. Representing the front, top and side views in the same proportions (Yue, 2006), these figures are easy to draw but difficult to understand for respondents who are not familiar with them (Yue, 2006). The MRT presents sets of cubes bent twice, characterised by salient features, such as the number of cubes and the orientation of the bends.

These characteristics are affordances, i.e. invitations to action, such as manipulation: “A rigid object with a sharp dihedral angle, an edge, affords cutting and scraping: it is a knife” (Gibson, 2015, p. 125). Drawing on Gestalt theory (Wertheimer, 1923), Gibson (2015) defined affordance as invariants, which observers may or may not perceive. It concerns objects as well as the environments where they are found. External to respondents and their needs, affordances are constantly available for perception, even if they are not always perceived. They are part of a “process of perceiving a rich-value ecological object” (Gibson, 2015, p. 132) and are “specified in stimulus information” (Gibson, 2015, p. 131). In the case of tests, Gibson’s definition of affordance was supplemented by Salmani Nodoushan (2021), who contrasted it with that of a test’s function, which concerns the specific and limited use of a test, whereas affordances encompass all subsidiary functions and uses. Like a chair, the main function of which is to allow someone to sit down but which can also be climbed on or thrown, tests have peripheral affordances that are different from the skill they are intended to measure. Hence, we propose that the characteristics of spatial tests, such as the nature of the stimuli or the method of administration, are affordances, because they structure the respondents’ mode of reasoning. For example, multiple-choice questions and the characteristics of the objects presented afford resolution strategies.

### ***Strategies for solving spatial tasks***

While these tests assess the ability to solve spatial tasks, they do not provide information about the activity engaged by the respondents. Part of the research dedicated to spatial ability focuses on the sources of variance in performance measured through spatial tests, including the resolution strategies implemented in the tasks presented (Eliot, 1987). According to Bloom & Broder (1950), because tests are designed to measure specific skills, the accurate interpretation of individual performance cannot be limited to the results. It must also “investigate the nature of the mental processes underlying the responses to these instruments” (p. 3). Similarly, Leplat (2004) considered that skill or competence is an abstract concept that can only be analysed through observable manifestations of activity. He posited that a task corresponds to a goal to be achieved and the conditions under which it must be achieved, while an activity concerns what is done by the subject to perform the task. This can be likened to the strategy described in the online APA Dictionary of Psychology (2018d)

as “a program of action designed to achieve a goal or accomplish a task”. Whether discussing activity or strategy, researchers can study task execution by observing respondents under normal conditions or under conditions created by the analyst, for example, through verbal protocols conducted during (Lemaire, 2015) or after the activity (Leplat, 2004). For example, Hegarty (2018) combined concurrent and retrospective interviews to examine the strategies adopted by 47 students for solving the questions in the second part of the MRT. Albaret & Aubert (1996) invited 288 high school students to describe in one sentence the method they used to answer the MRT questions. These two studies revealed the use of different strategies to solve a test that aims to measure mental rotation, including mental rotation, comparing the stimuli using a salient element (considered an analytical strategy (Eme & Marquer, 1998)), and the combination of several strategies. However, these observations depend on the participants’ ability to verbalise mental processes (Bloom & Broder, 1950; Pilardeau, 2008), of which they are not always aware (Bloom & Broder, 1950; Lemaire, 2015; Lohman, 1979).

Another solution is to infer strategies from performance indicators, such as completion time. For example, Lohman and Kyllonen (1983) assigned completion times to previously defined strategy models. However, according to Lemaire (2015), indirect investigation methods risk confusing different strategies achieved in identical response times. He therefore recommended studying strategies directly, whenever possible, to minimise interpretations and inferences. According to Dunlosky and Hertzog (2001), retrospective reports are suitable for describing problem-solving strategies when they take place immediately after the test, to reduce the risk of forgetting.

These studies revealed that spatial tests, designed to assess a specific skill, can mobilise strategies other than the one targeted by the test. Other studies revealed that participants use several strategies in the same test (Hegarty, 2018; Workman et al., 1999), and that some change strategies from one question to another (Pilardeau, 2008; Workman et al., 1999).

Our initial literature review aimed to characterise the relationship between spatial skills and 3-D modelling (Charles, 2023). Currently, this field of research is mainly investigated by authors in engineering sciences (e.g. Martín-Gutiérrez et al., 2013; Metz et al., 2016), who seek to design and evaluate educational tools for remedying spatial ability. This

educational support is considered useful for increasing student retention in science programmes (Sorby & Veurink, 2010), as there is a confirmed correlation between spatial performance and success in STEM (Wai et al., 2009). This approach focuses on spatial tests as measures of competence and does not address problem-solving strategies, explaining why these studies have not addressed the issue of strategies, unlike research conducted in the field of psychology (Hegarty, 2018; Hegarty & Waller, 2004).

## Methodology

Our methodology was developed in three phases of experimentations and results: (a) identification and use of the most widely employed tests in the study of the link between spatial ability and volume modelling (Experiment A); (b) investigation into the activity of respondents taking part in spatial tests (Experiment B); and, (c) the final methodology, which combined spatial tests and observation of the activity involved in these measurements (Experiment C).

### *Experiment A*

The first step in this process was to investigate spatial tests. This involved identifying, obtaining, and experimenting with tests. The most readily available tests were those most cited in the literature review addressing the link between spatial ability and engineering studies (Ault & John, 2010; Kelly Jr, 2013): the MRT, the MCT, and the original version of the R PSVT:R, the Purdue Spatial Visualisation Test: Rotations (Guay, 1976b). We then found and selected Yoon's revised version (2011) as it is not time-limited and therefore potentially less disadvantageous to women, as mentioned by Cooke-Simpson and Voyer (2007).

In September 2018, 137 first-year engineering students,  $N_w = 37$  [27%] women and  $N_M = 100$  [73%] men, took the MRT and MCT, and 131 students,  $N_w = 36$  [27%] females and  $N_M = 95$  [73%] males, took the R PSVT:R, according to the authors' administration instructions. The scores were compared with the marks obtained in Applied Mathematics (MAPP), which addresses the theory of distributions of a real variable and the Laplace transform, in Mechanism Analysis (AMEC), which aims to teach students to analyse an industrial mechanism, and in Algorithms and Programming (ALGO), which addresses algorithms, fundamental

data structures, and the C language. An analysis of variance (ANOVA) highlighted the predictive nature of the MCT, MRT, and R PSVT:R of performance in AMEC, ALGO, and MAPP (Charles et al., 2019).

### ***Experiment B***

Adopting an exploratory approach, the first five questions from the tests obtained were given to different types of respondents. For the first experiments, we sampled people from our personal network ( $N_C = 2$  children and  $N_A = 2$  adults) and filmed them in problem-solving situations. They revealed differences in behaviour: Some used spontaneous hand or head gestures, as observed by Chu and Kita (2008). These observations were then supplemented by retrospective interviews to question the individuals' mental activity. This aimed to identify the strategies used in the collected tests for the subsequent selection of four tests that measure each of the four factors in Tartre's model, illustrated in Figure 1. This protocol was extended to a sample of children volunteers from our personal network ( $n = 4$ ), first-year engineering students from the school where the main data collection would take place ( $n = 8$ ), students in sports science and technology (STAPS) attending a partner institution of our university ( $n = 6$ ), and professionals from our personal network ( $n = 4$ ). This broad sample was chosen to avoid collecting problem-solving strategies specific to a given age or specialisation. Some of the resolution processes described by the respondents for the same test differed from one person to another, or from one question to another. Table 1 shows the frequency of strategies reported as used in the spatial tests listed below:

- Mental rotation tests: the MRT and the R PSVT:R;
- Mental transformation tests: the MCT and the Santa Barbara Solids Test (SBST) (Cohen and Hegarty, 2007);
- Perspective change tests: the PSVT:V, Visualization of Viewpoints (VVT) (Guay & McDaniel, 1976), and the Perspective Taking/ Spatial Orientation Test (PTSOT) (Hegarty & Waller, 2004; Kozhevnikov & Hegarty, 2001);
- Disembedding tests: the CFT, the Hidden Figures Test (HFT) (Ekstrom et al., 1976), and the Hidden Patterns Test (HPT) (Ekstrom et al., 1976).

Table 1 shows that these tests engage different or multiple solution strategies, including strategies that differ from those targeted.

Table 1  
*Frequency of strategies reported as used in Experiment B (N = 22)*

Test	n	Strategy used												
		Mental transformation	Mental rotation	Spatial orientation	Analytical strategy	Physical strategy	MR/SO combination	MR/AS combination	SO/AS combination	MR/SO/AS combination	MR/MT combination	MT/AS combination	RM/PS combination	MR/AS combination
VVT	14	0	3	3	4	0	1	0	2	0	0	0	1	0
PTSOT	14	0	2	8	0	1	3	0	0	0	0	0	0	0
R PSVT:R	20	0	<b>18</b>	0	0	0	0	0	0	1	1	0	0	0
PSVT:V	17	0	0	<b>14</b>	0	0	2	0	1	0	0	0	0	0
MRT	9	0	5	0	0	0	0	4	0	0	0	0	0	0
MCT	15	7	0	0	0	0	0	0	0	0	0	7	0	1
CFT	13	0	0	3	<b>6</b>	0	1	0	3	0	0	0	0	0
HFT	11	0	3	1	3	0	1	0	3	0	0	0	0	0
SBST	4	4	0	0	0	0	0	0	0	0	0	0	0	0
HPT	2	0	0	0	2	0	0	0	0	0	0	0	0	0

*Note.* mental rotation (Tartre, 1984), MT = mental transformation (Tartre, 1984), SO = spatial orientation (Tartre, 1984), AS = analytical strategy (Eme & Marquer, 1998), PS = physical strategy (Chu & Kita, 2008).

### *Experiment C*

Following Experiment B, we selected the tests that most effectively measured one of the four factors in Tartre's model (1984): the R PSVT:R for mental rotation, the PSVT:V for perspective change, the MCT for mental transformation, and the CFT for disembedding. The MRT was added because of its predictive value of engineering students' academic success identified in Experiment A, namely in the AMEC assessment which requires the use of a 3-D modeller, but also for comparative purposes with the numerous studies on spatial skills and volume modelling (Ault & John, 2010; Kelly Jr, 2013).

In light of the characteristics of spatial tests, as noted in the literature and observed in Experiments A and B, we determined a methodology for investigating spatial abilities to report on the skills used in spatial tests: measures of the students' performances in spatial tasks and investigations of the cognitive processes involved. To that end, we supplemented the

spatial tests, which collect quantitative data on participants' ability to solve spatial problems, with retroactive reports to qualitatively control the object of these measurements.

### *Participants*

The experiment used a convenience sample of 146 first-year engineering students, comprising  $N_w = 27$  [18.5%] women and  $N_m = 119$  [81.5%] men, from a variety of academic backgrounds: scientific preparatory classes for French elite higher education institutions, bachelor's degree programmes, advanced technician programmes, and university training institutes. A total of 80 [55%] of the cohort had received technology instruction in the second year of these programmes.

### *Tools for measuring spatial ability*

Five spatial tests were selected to assess Tartre's four spatial factors (1984): The R PSVT:R and the MRT for assessing mental rotation, the MCT for mental transformation, the PSVT:V for the ability to change perspective, and the CFT for the ability to isolate an element embedded in a complex pattern.

We followed the administration guidelines available in the test instructions for the CFT. We used Eliot and Macfarlane Smith's guidelines (1983) for the PSVT:V and the MCT, and the administration times commonly cited in the literature for the MRT, i.e. three minutes for the first part, a two-minute break, and three minutes for the second part, after receiving confirmation from one of the authors that these times were appropriate for our sample (A. R. Kuse, personal communication, 25 June 2018). We limited the administration time for the R PSVT:R to one hour for scheduling reasons. As Experiment B confirmed the use of alternative strategies, we placed the tests shown to be most sensitive to circumvention strategies at the beginning of the battery for Experiment C.

### *Investigation into the strategies used*

Due to the number of participants and the time available in the academic calendar, we conducted consecutive observations (Leplat, 2004): Spatial tests were followed by open and closed questionnaires to explore the problem-solving processes used. First, we asked students to explain the process they had followed to answer three questions from each test. In each case, they were asked to mentally find the solution and then describe the resolution process. The aim was to gather the most recent and

detailed possible recollection of the strategy used, such as the nature and order of the steps, errors, and dead ends (Bloom & Broder, 1950). Three questions per test were chosen to avoid a response that could apply to all questions in the same test, as is the case for the MCQ described in the following paragraph. We selected items presenting different types of difficulties: The number of sides (Kyllonen et al., 1981), the nature of surfaces (Pellegrino et al., 1985), the amplitude and number of rotations (Cooper & Mumaw, 1985; Guay, 1980), the number of elements composing a stimulus (Pellegrino et al., 1985), and the presence of occlusions (Caissie et al., 2009; Pellegrino et al., 1985). The questions at the end of the test were discarded in case not all of the students had time to answer them.

The respondents were then asked to choose the strategies they thought they had used in a test from a multiple-choice list. These statements were selected from the strategies most frequently described by the engineering students ( $n = 8$ ) and sports students ( $n = 6$ ) tested in Experiment B. Students could choose as many answers as they wished. The MCQ complemented the answers given in the open-ended questionnaire (OEQ), as it focuses on strategies used in a test, while the OEQ explores strategies specific to three questions per test. The OEQ was conducted before the MCQ to prevent the answers proposed in the MCQ from influencing the students' answers to the OEQ (Schwarz & Hippler, 1991).

### ***Ethical considerations***

Students were informed of the framework and objective of the study, as well as the procedure. In accordance with French Law 2018-493 of 20 June 2018, they were informed of their right to access the data collected during the experiment. They signed a consent form previously approved by the school's data protection officer. All data were anonymised.

### ***Methods of analysis***

Scores were calculated using the instructions available in the manual for the CFT (Thurstone & Jeffrey, 1965), in *An International Directory of Spatial Tests* (Eliot & Macfarlane Smith, 1983) for the MCT and PSVT:V, and according to the authors' recommendations for the R PSVT:R and MRT (Maeda et al., 2013; Vandenberg & Kuse, 1978).

The verbatim responses collected in the OEQ were encoded following strategies identified in the literature review, noting overall strategies, such as mental rotation, and the steps therein, for example (a) determining the identity and orientation of the object; (b) rotating the object; (c)

comparing it with its usual representation; and (d) responding (Cooper & Shepard, 1973). The constituent steps of a global strategy were grouped into global strategies whenever possible. Observables of mode of explanation (e.g., drawing that expresses response), mode of response (e.g., response identical to the previous response indicated by ‘idem’), sequences indicated by sequencing markers (e.g., then), and articulation of reflection (e.g., I deduce) were also noted. The coding kit, submitted to two external observers, was validated with agreement rates between 96% and 98%. See our doctoral dissertation (Charles, 2023) for a detailed explanation of this methodology. The answers chosen in the MCQ were coded according to the overall strategies mentioned in the literature review (Chu & Kita, 2008; Eme & Marquer, 1998; Tartre, 1984).

We encoded the verbatim responses to calculate the number of different strategies reported as used in a test for the three chosen questions and then determined whether the students had changed strategies between at least two of the chosen questions. Students who used a generic response to all three questions, rather than a specific response to each question, were excluded.

## Results

### *Spatial performance*

Here we present some of the data from our doctoral findings: The distribution of scores varies according to the tests, including the two tests that target the same spatial skill, i.e. the MRT and the R PSVT:R. As Table 2 shows, the distributions of scores for the PSVT:V, R PSVT:R, MRT, and MCT do not follow a normal distribution:

- PSVT:V: dispersion of low and average scores, while the distribution of scores between 25 and 30 appears normal;
- R PSVT:R: dispersion of low and average scores, while the distribution of scores between 24 and 30 appears normal;
- MRT: dispersion of low scores, while the distribution of scores between 8 and 20 appears normal;
- MCT: dispersion of low scores, while the distribution of scores between 9 and 25 appears normal.

Table 2  
*Descriptive statistics and normality of spatial scores (N = 146)*

Test	<i>N</i>	Min	Max	<i>M</i>	<i>SD</i>	<i>W</i>	<i>p</i>
PSVT:V	144	5	30	25.10	5.56	0.75	< 0.001
R PSVT:R	146	8	30	25.31	4.06	0.87	< 0.001
MRT	146	0	20	12.55	4.42	0.96	< 0.001
MCT	146	3	25	16.16	5.13	0.96	< 0.001
CFT	146	22	160	96.52	26.96	0.99	NS

*Note.* *N* = number of students present for the tests; *M* = mean; *SD* = standard deviation; *W* = Shapiro-Wilk test statistic; *p* = p-value; NS = not significant.

The majority of students obtained average and high scores, except for the CFT where distribution was normal.

### *Resolution strategies adopted in the spatial tests*

#### *Strategies identified in the open-ended questionnaire*

Encoding the verbatim responses to the first question chosen from each test determined the number of students who reported using the skill targeted by the test, and the number of students who described using a strategy other than that targeted by the test for that question. Table 3 shows the percentage of students who stated they used a strategy other than that targeted by the test, ranging from 6% to 65% depending on the test.

Table 3  
*Number of students who stated they used the strategies targeted in the tests*

Test	<i>N</i>	Targeted skill	Skill used			
			Skill targeted by the test		Other strategy	
			<i>n</i>	%	<i>n</i>	%
PSVT:V	145	Perspective change	111	77	34	23
R PSVT:R	144	Mental rotation	135	94	9	6
MRT	140	Mental rotation	90	64	50	36
MCT	142	Mental transformation	73	51	69	49
CFT	142	Disembedding	50	35	92	65

*Note.* *N* = number of students whose responses were used; *n* = number of students who reported using the strategy; % = percentage of students who recalled using the strategy

Regarding the PSVT:V, which aims to measure the ability to change a point of view,  $N_{MR} = 15$  [10%] students reported using mental rotation and  $N_C = 19$  [13%] combinations of several strategies in the first question selected. For the R PSVT:R, which intends to measure mental rotation,  $N_{PS} = 2$  [1%] students described the change of perspective strategy, and  $N_C = 7$  [5%] reported a combination of strategies for solving the first selected question. Regarding the first question selected from the R PSVT:R, which targets mental rotation,  $N_{PS} = 2$  [1%] students recalled a change of perspective and  $N_C = 7$  [5%] indicated a combination of strategies. For the first question selected from the MRT, which intends to measure mental rotation,  $N_{US} = 32$  [23%] students described a single strategy, such as changing their point of view or comparing using a salient element, and  $N_C = 18$  [13%] reported a combination of strategies. Regarding the first question chosen from the MCT, which seeks to measure mental transformation,  $N_{US} = 41$  [29%] students recalled single strategies, such as mathematical knowledge or a real-life situation, and  $N_C = 28$  [20%] indicated a combination of strategies. Finally, for the first question chosen from the CFT, which aims to measure disembedding,  $N_{US} = 62$  [44%] students described single strategies such as mental rotation or imagining themselves drawing, and  $N_C = 30$  [21%] reported a combination of strategies.

Verbatim encoding was used to compare the strategies used in three questions for each test in the battery and to determine whether students changed strategies from one question to another. Table 4 shows that between 32% and 62% of students reported multiple strategies depending on the test.

Table 4  
*Number of students reporting the use of multiple strategies in verbatim responses and multiple-choice questions*

Test	Open-ended questionnaire			Multiple-choice questionnaire		
	<i>N</i>	<i>n</i>	%	<i>N</i>	<i>n</i>	%
PSVT:V	143	76	53	144	98	67
R PSVT:R	142	45	32	146	51	35
MRT	138	46	33	142	84	60
MCT	125	77	62	146	85	58
CFT	136	51	37,5	146	52	35,5

*Note.* *N* = total number of students whose responses were used; *n* = number of students who reported using multiple strategies; % = percentage of students who recalled using multiple strategies.

The descriptive statistics for the number of strategies reported as used in the five tests in Table 5 indicate that students used an average of one or two strategies per test.

### *Strategies identified in the multiple-choice questions*

Encoding the answers chosen in the MCQ makes it possible to identify the strategies used for each test in our battery, whether the strategy was used alone, such as mental rotation, or part of a combination of strategies, such as mental rotation and disembedding. This calculation does not determine whether the strategies are used in isolation or whether they are combined to answer a question. For example, if the student answered mental rotation and disembedding, it is not possible to determine whether they used mental rotation for some of the questions and disembedding for others, or whether they used a combination of these two strategies for all the questions or only for some of them. We thus observe that a percentage of students, ranging from 35% to 67% depending on the test, recalled several unique strategies or a combination of strategies (Table 4). The descriptive statistics for the number of strategies reported in the five tests in Table 5 indicate that students used either one or two unique strategies or combinations of strategies per test on average.

Table 5

*Descriptive statistics for the number of strategies reported as used in the verbatim and multiple-choice questionnaire*

Test	Open-ended questionnaire						Multiple-choice questionnaire					
	<i>N</i>	<i>M</i>	Mdn	<i>SD</i>	Min	Max	<i>N</i>	<i>M</i>	Mdn	<i>SD</i>	Min	Max
PSVT:V	143	1.61	2	0.628	1	3	144	1.8	2	0.65	1	3
R PSVT:R	142	1.35	1	0.533	1	3	146	1.4	1	0.593	1	4
MRT	138	1.39	1	0.597	1	3	146	1.65	2	0.586	1	3
MCT	125	1.83	2	0.759	1	3	146	1.64	2	0.596	1	3
CFT	136	1.43	1	0.605	1	3	142	1.44	1	0.643	1	3

*Note.* *N* = number of students whose responses were used; Min = lowest number of strategies observed; Max = highest number of strategies observed; *M* = mean number of strategies observed; *SD* = measure of dispersion of values; Mdn = median.

## Discussion

The scores obtained by our sample describe a performance that is mostly average to high for four tests in our battery. This result can be explained by the fact that our respondents are students specialised in

STEM subjects, i.e., students who are likely to have high spatial skills (Wai et al., 2009). However, this observation does not apply to the results observed for the CFT, the disembedding test. We propose our sample did not follow academic studies that implement learning situations conducive to the development of this skill, which some students may have acquired through certain individual practices, such as leisure activities.

Table 4 shows that some students used several strategies for all the tests in our battery, whether single strategies or combinations of strategies, and that students changed strategies from one question to another. Whether single or combined, the strategies used did not always correspond to the skills targeted by the test. For example, some students rotated objects rather than changing their point of view to indicate what an object looked like when viewed from an angle different from that presented in the stimulus in the PSVT:V, a test designed to measure changes in point of view. In the CFT, a disembedding test, some students used a combination of strategies to isolate an element in a pattern: They started by moving the stimulus into the complex pattern, then searched for it in the pattern. However, we note that in the case of the R PSVT:R, a mental rotation test, 94% of students indicated using mental rotation alone in the MCQ, and that 90 to 94% reported using mental rotation alone for the three selected questions.

The results of our study confirm those observed in the literature (Albaret & Aubert, 1996; Hegarty, 2018; Pilardeau, 2008) for first-year engineering students. They point to the fact that performance on spatial tests cannot be solely explained by proficiency in the skills targeted by the tests. To precisely identify the skills actually being measured, performance measurements must be complemented with protocols for investigating strategies used in spatial tests. Spatial performance, as measured by these tests, reveals proficiency in the skill of solving problems of a spatial nature, rather than proficiency in a specific spatial skill. Caution is therefore recommended when interpreting the skills supposedly measured. It is important to question the relationship between these determining skills and performance in pursuing STEM studies (Wai et al., 2009), whether the results are used for selection purposes (Maeda & Yoon, 2011) or for designing and/or validating courses for developing these skills (Martín-Gutiérrez et al., 2015).

### *Limitations and perspectives*

The order of the tests in the battery was the same for all students, as only one session was organised and the tests were limited in time. As a result, the strategies and performances observed in the tests administered after the first test may have been affected by the order effect (American Psychological Association, 2018a; Kinnear & Gray, 2015): incidental learning in the first tests may have influenced the following tests.

The activity analysis methodology we adopted is limited to students' ability to represent and explain the cognitive processes involved in their action (Bloom & Broder, 1950; Lemaire, 2015; Pilardeau, 2008). The strategies identified in the verbatim were also limited to the three questions chosen for each test. In addition, the list of strategies proposed in the MCQ may have influenced the students' choices of answers (Schwarz & Hippler, 1991).

As the validity of retrospective protocols has been established for people who do not suffer from memory disorders (Dunlosky & Hertzog, 2001), we plan to use the strategies identified in our work to design an MCQ to identify the cognitive processes at work in spatial tests. Despite the limitations of these tools, it is important to highlight the skills used in spatial performance measurement tools. Given the selected items for the open questionnaire were difficult for different reasons, such as the number of sides (Kyllonen et al., 1981) or the nature of the surfaces (Pellegrino et al., 1985), it would also be interesting to investigate the relationship between the strategies used and performance by item to explore whether some strategies were more effective than others for solving specific items.

### **Conclusion**

This article aimed to present a methodology for characterising the spatial abilities of first-year engineering students, developed as part of the e-FRAN EXAPP\_3D project. This approach found that, although spatial tests are designed to assess a specific skill, they do not always succeed in engaging the skill they aim to assess, in other words, they do not sufficiently afford these skills. This seems to be less true for the R PSVT:R, which has two particularities: No time limit, and problems for which it is not possible to simply compare the stimuli with the possible answers.

Affordance due to the need to use mental rotation on an example, before applying it to a stimulus, seems sufficiently robust to avoid circumvention strategies, and to produce a performance that reveals the targeted skill. The affordance of the question seems to outweigh that of the absence of a time limit or, in our case, the allocation of generous response time, which allows for the implementation of an increased number of strategies (Schwartz, 1963). Our results suggest that it may be pertinent to reconsider the use of these tests for assessing spatial skills, as they may confuse performance with competence.

Proofreading: Caroline Lefour

Formatting: Emmanuel Gagnon

Portuguese abstract: Eusebio Andre Machado

## REFERENCES

- Albaret, J. M., & Aubert, E. (1996). Étalonnage 15-19 ans du test de rotation mentale de Vandenberg. *EVOLUTIONS psychomotrices*, 8(34), 269–278.
- American Psychological Association. (2018a). Order effect. In *APA Dictionary of Psychology*. <https://dictionary.apa.org/order-effect>
- American Psychological Association. (2018b). Performance. In *APA Dictionary of Psychology*. <https://dictionary.apa.org/performance>
- American Psychological Association. (2018c). Performance test. In *APA Dictionary of Psychology*. <https://dictionary.apa.org/performance-test>
- American Psychological Association. (2018d). Strategy. In *APA Dictionary of Psychology*. <https://dictionary.apa.org/strategy>
- Ault, H. K., & John, S. (2010). Assessing and Enhancing Visualisation Skills of Engineering Students in Africa: A Comparative Study. *Engineering Design Graphics Journal*, 74(2), 12–20.
- Barratt, E. S. (1953). An analysis of verbal reports of solving spatial problems as an aid in defining spatial factors. *The Journal of Psychology*, 36, 17–25.
- Bloom, B. S., & Broder, L. J. (1950). Problem-solving processes of college students. *Supplementary Educational Monographs*, B, 1–109.
- Branoff, T. (2000). Spatial Visualisation Measurement: A Modification of the Purdue Spatial Visualisation Test - Visualisation of Rotations. *Engineering Design Graphics Journal*, 64(2), 14–22.
- Branoff, T., & Dobelis, M. (2012). The Relationship between Spatial Visualisation Ability and Students' Ability to Model 3D Objects from Engineering Assembly Drawings. *Engineering Design Graphics Journal*, 76(3), 37–43.
- Caissie, A. F., Vigneau, F., & Bors, D. A. (2009). What does the Mental Rotation Test Measure? An Analysis of Item Difficulty and Item Characteristics. *The Open Psychology Journal*, 2(1), 94–102. <https://doi.org/10.2174/1874350100902010094>
- Charles, S. (2023). *Habilité spatiale et stratégies de modélisation 3D* (Doctoral dissertation, CY Cergy Paris Université, Cergy, France). <https://hal.science/tel-04097396>
- Charles, S., Jaillet, A., Peyret, N., Jeannin, L., & Rivière, A. (2019). Exploring the relationship between spatial ability, individual characteristics and academic performance of first-year students in a French engineering school. *47th SEFI Annual Conference Proceedings*, 235–248. [https://www.sefi.be/wp-content/uploads/2019/10/SEFI2019\\_Proceedings.pdf](https://www.sefi.be/wp-content/uploads/2019/10/SEFI2019_Proceedings.pdf)
- Chu, M., & Kita, S. (2008). Spontaneous gestures during mental rotation tasks: Insights into the microdevelopment of the motor strategy. *Journal of Experimental Psychology: General*, 137(4), 706–723. <https://doi.org/10.1037/a0013157>
- Cohen, C. A., & Hegarty, M. (2007). Sources of Difficulty in Imagining Cross Sections of 3D Objects. *Proceedings of the 29th Annual Meeting of the Cognitive Science Society*, 179–184.
- College Entrance Examination Board. (1939). *Special Aptitude Test in Spatial Relations (Mental Cutting Test)*. College Entrance Examination Board.

- Cooke-Simpson, A., & Voyer, D. (2007). Confidence and gender differences on the Mental Rotations Test. *Learning and Individual Differences, 17*(2), 181–186. <https://doi.org/10.1016/j.lindif.2007.03.009>
- Cooper, L. A., & Mumaw, R. J. (1985). Spatial Aptitude. In R. F. Dillon (Ed.), *Individual Differences in Cognition* (Vol. 2, pp. 67–94). Academic Press.
- Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual Information Processing* (pp. 75–176). Academic Press. <https://doi.org/10.1016/B978-0-12-170150-5.50009-3>
- Dunlosky, J., & Hertzog, H. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition, 29*(2), 247–253.
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Educational Testing Service.
- Eliot, J. (1983). Historical background. In *An international directory of spatial tests* (pp. 1–10). NFER-Nelson.
- Eliot, J. (1987). *Models of Psychological Space: Psychometric, Developmental, and Experimental Approaches*. Springer-Verlag.
- Eliot, J., & Macfarlane Smith, I. (1983). *An international directory of spatial tests*. NFER-Nelson.
- Eme, P. E., & Marquer, J. (1998). Quantitative and qualitative individual differences in visual comparison processing: Are strategies only analytic or holistic? *Personality and Individual Differences, 25*(3), 459–475. [https://doi.org/10.1016/S0191-8869\(98\)00049-X](https://doi.org/10.1016/S0191-8869(98)00049-X)
- French, J. W. (n.d.). The relationship of problem-solving styles to the factor composition of tests. *Educational and Psychological Measurement, 25*(1), 9–28.
- Galton, F. (1883). *Inquiries into human faculty and its development* (G. Tredoux, Ed.; 2001 1st electronic ed.). <http://galton.org/>. <http://galton.org/books/human-faculty/text/human-faculty.pdf>
- Geronimi, A., De Vries, E., Prudhomme, G., & Baille, J. (2005). Objets intermédiaires dans une situation de conception en technologie avec CAO au collège. *Aster, 41*, 115–137. <https://doi.org/10.4267/2042/8869>
- Gibson, J. J. (2015). *The Ecological Approach to Visual Perception* (Classic Edition). Psychology Press.
- Guay, R. B. (1976a). *Purdue Spatial Visualisation Test: Rotations*. Purdue Research Foundation.
- Guay, R. B. (1976b). *Purdue Spatial Visualisation Test—Visualisation of Views*. Purdue Research Foundation.
- Guay, R. B. (1980, April 7). *Spatial Ability Measurement: A Critique and an Alternative* (Communication paper). Annual Meeting of the American Educational Research Association, Boston, MA.
- Guay, R. B., & McDaniel, E. D. (1976). *The visualisation of viewpoints*. Purdue Research Foundation.
- Guay, R. B., McDaniel, E. D., Angelo, S., Guay, R. B., & McDaniel, E. (1978). Analytic factor confounding spatial ability measurement. In *Correlates of performance on spatial aptitude tests*. (pp. 116–128). U.S. Army Research Institute for the Behavioral and Social Sciences.

- Hamon, C. (2009). Graphismes techniques: Tâches, nature et causes des difficultés des apprenants. *Aster*, 48, 39–62. <https://doi.org/10.4267/2042/30420>
- Hegarty, M. (2010). Components of Spatial Intelligence. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 52, pp. 265–297). Academic Press. [https://doi.org/10.1016/S0079-7421\(10\)52007-3](https://doi.org/10.1016/S0079-7421(10)52007-3)
- Hegarty, M. (2018). Ability and sex differences in spatial thinking: What does the mental rotation test really measure? *Psychonomic Bulletin & Review*, 25, 1212–1219.
- Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32(2), 175–191.
- Hopkins, K. D. (1998). *Educational and Psychological Measurement and Evaluation* (8th ed.). Allyn & Bacon.
- Kelly Jr, W. F. (2013). *Measurement of Spatial Ability in an Introductory Graphic Communications Course* (Doctoral dissertation, North Carolina State University).
- Kinncar, P., & Gray, C. (2015). *SPSS facile appliqué à la psychologie et aux sciences sociales: Maîtriser le traitement des données*. De Boeck.
- Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Memory & Cognition*, 29(5), 745–756. <https://doi.org/10.3758/BF03200477>
- Kyllonen, P. C., Woltz, D. J., & Lohman, D. F. (1981). *Models of Strategy and Strategy-Shifting in Spatial Visualisation Performance*. (Technical report No. 17). School of Education, Stanford University. <https://apps.dtic.mil/sti/citations/ADA108003>
- Leclercq, D. (2006). *L'évolution des QCM*. In G. Figari & L. Mottier-Lopez (Eds), *Recherches sur l'évaluation en Education* (pp. 139–146). L'Harmattan. <https://orbi.uliege.be/handle/2268/10124>
- Lemaire, P. (2015). *Viellissement cognitif et adaptations stratégiques* (1st ed.). De Boeck Supérieur.
- Leplat, J. (2004). L'analyse psychologique du travail. *European Review of Applied Psychology*, 54(2), 101–108. <https://doi.org/10.1016/j.erap.2003.12.006>
- Lieu, D. K., & Sorby, S. A. (2009). *Visualisation, modeling, and graphics for engineering design* (1st ed). Delmar, Cengage Learning.
- Linn, M. C., & Petersen, A. C. (1985). Emergence and Characterization of Sex Differences in Spatial Ability: A Meta-Analysis. *Child Development*, 56(6), 1479–1498. JSTOR. <https://doi.org/10.2307/1130467>
- Lohman, D. F. (1979). *Spatial Ability: A Review and Reanalysis of the Correlational Literature*. (Technical report No. 8). Stanford University, School of Education. <https://apps.dtic.mil/sti/citations/ADA075972>
- Lohman, D. F. (1993). Spatial Ability and G. In I. Dennis & P. Tapsfield (Eds), *Human Abilities: Their Nature and Measurement* (1st ed, pp. 97–116). Psychology Press.
- Lohman, D. F., & Kyllonen, P. C. (1983). Individual Differences in Solution Strategy on Spatial Tasks. In R. F. Dillon & R. R. Schmeck (Eds). In *Individual Differences in Cognition* (Vol. 1, pp. 105–135). Academic Press.
- Lohman, D. F., Pellegrino, J. W., Alderton, D. L., & Regian, J. W. (1987). Dimensions and Components of Individual Differences in Spatial Abilities. In S. H. Irvine & S. E. Newstead (Eds), *Intelligence and Cognition: Contemporary Frames of Reference* (pp. 253–312). Springer Netherlands. [https://doi.org/10.1007/978-94-010-9437-5\\_6](https://doi.org/10.1007/978-94-010-9437-5_6)

- Maeda, Y., & Yoon, S. Y. (2011). Scaling the Revised PSVT-R: Characteristics of the first-year engineering students' spatial ability. *ASEE Annual Conference and Exposition, Conference Proceedings*. <https://doi.org/10.13140/2.1.2600.2881>
- Maeda, Y., Yoon, S. Y., Kim-Kang, G., & Imbrie, P. K. (2013). Psychometric Properties of the Revised PSVT-R for Measuring First Year Engineering Students' Spatial Ability. *International Journal of Engineering Education*, 29(3), 763–776.
- Maier, P. H. (1996). Spatial geometry and spatial ability—How to make solid geometry solid. *Selected Papers from the Annual Conference of Didactics of Mathematics*, 63–75.
- Martín-Gutiérrez, J., Contero, M., & Alcañiz, M. (2015). Augmented Reality to Training Spatial Skills. *Procedia Computer Science*, 77, 33–39. <https://doi.org/10.1016/j.procs.2015.12.356>
- Martín-Gutiérrez, J., Gil, F. A., Contero, M., & Saorín, J. L. (2013). Dynamic three-dimensional illustrator for teaching descriptive geometry and training visualisation skills. *Computer Applications in Engineering Education*, 21(1), 8–25. <https://doi.org/10.1002/cae.20447>
- McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, 86(5), 889–918. <https://doi.org/10.1037/0033-2909.86.5.889>
- Metz, S. S., Sorby, S. A., & Jarosewich, T. (2016). Spatial Skills Training Impacts Retention of Engineering Students—Does This Success Translate to Community College Students in Technical Education? *2016 ASEE Annual Conference & Exposition*. DOI:10.18260/1-2--27719
- Necker, L. A. (1832). LXI. Observations on some remarkable optical phenomena seen in Switzerland; and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 1(5), 329–337. <https://doi.org/10.1080/14786443208647909>
- Onyancha, R. M., Derov, M., & Kinsey, B. L. (2009). Improvements in Spatial Ability as a Result of Targeted Training and Computer-Aided Design Software Use: Analyses of Object Geometries and Rotation Types. *Journal of Engineering Education*, 98(2), 157–167. <https://doi.org/10.1002/j.2168-9830.2009.tb01014.x>
- Pellegrino, J. W., Mumaw, R. J., & Shute, V. J. (1985). Analyses of spatial aptitude and expertise. In S. E. Embretson (Ed.), *Test design: Developments in psychology and psychometrics* (pp. 45–76). Academic Press.
- Pilardeau, M. (2008). *Identification et validation des stratégies dans une tâche de rotation d'image mentale* (Doctoral dissertation, Université Paris Descartes, Paris, France).
- Salmani Nodoushan, A. M. (2021). Test Affordances or Test Function? Did We Get Messick's Message Right? *International Journal of Language Studies*, 15(3), 127–140.
- Schwartz, P. A. (1963). Adapting tests to the cultural setting. *Educational and Psychological Measurement*, 23(4), 673–686.
- Schwarz, N., & Hippler, H.-J. (1990). *Response alternatives: The impact of their choice and presentation order* (ZUMA-Arbeitsbericht No. 1990/08). Zentrum für Umfragen, Methoden und Analysen. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-67257>
- Sorby, S. A., & Veurink, N. (2010). Long-term Results from Spatial Skills Intervention among First-year Engineering Students. *Proceedings of the 65th Midyear Meeting of the Engineering Design Graphics Division of ASEE*.

- Steinhauer, H. M. (2012). Correlation Between a Student's Performance on the Mental Cutting Test and Their 3D Parametric Modeling Ability. *Engineering Design Graphics Journal*, 76(3), 44–48.
- Tartre, L. A. (1984). *The Role of Spatial Orientation Skill in the Solution of Mathematics Problems and Associated Sex-Related Differences*. (Doctoral dissertation, The University of Wisconsin-Madison).
- Tartre, L. A. (1990). Spatial skills, gender, and mathematics. In E. Fennema & G. C. Leder (Eds), *Mathematics and gender* (pp. 27–59). Teachers College Press.
- Thurstone, L. L. (1938). The perceptual factor. *Psychometrika*, 3(1), 1–17. <https://doi.org/10.1007/BF02287914>
- Thurstone, L. L., & Jeffrey, T. E. (1956). *Closure Flexibility (Concealed Figures) Test—Form A*. Industrial Relations Center - The University of Chicago.
- Thurstone, L. L., & Jeffrey, T. E. (1965). *Closure Flexibility (Concealed figures) Test administration Manual*. Industrial Relations Center - The University of Chicago.
- Tsutsumi, E., Shiina, K., Suzaki, A., Yamanouchi, K., Saito, T., & Suzuki, K. (1999). A mental cutting test on female students using a stereographic system. *Journal for Geometry and Graphics*, 3(1), 111–119.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. *Psychological Bulletin*, 139(2), 352–402. <https://doi.org/10.1037/a0028446>
- Vandenberg, S. G., & Kuse, A. R. (1978). Mental Rotations, a Group Test of Three-Dimensional Spatial Visualisation. *Perceptual and Motor Skills*, 47(2), 599–604. <https://doi.org/10.2466/pms.1978.47.2.599>
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial Ability for STEM Domains: Aligning over 50 Years of Cumulative Psychological Knowledge Solidifies Its Importance. *Journal of Educational Psychology*, 101(4), 817–835. <https://doi.org/10.1037/a0016127>
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung*, 4(1), 301–350. <https://doi.org/10.1007/BF00410640>
- Workman, J. E., Caldwell, L. F., & Kallal, M. J. (1999). Development of a Test to Measure Spatial Abilities Associated with Apparel Design and Product Development. *Clothing and Textiles Research Journal*, 17(3), 128–133. <https://doi.org/10.1177/0887302X9901700303>
- Yoon, S. Y. (2011). *Psychometric properties of the Revised Purdue Spatial Visualisation Tests: Visualisation of Rotations (the Revised PSVT:R)* (Doctoral dissertation, Purdue University). <https://search.proquest.com/openview/e5c5ec5adee291928859d5158fa37e201/pq-origsite=gscholar&cbl=18750&diss=y>
- Yoon, S. Y., & Mann, E. L. (2017). Exploring the Spatial Ability of Undergraduate Students: Association with Gender, STEM Majors, and Gifted Program Membership. *Gifted Child Quarterly*, 61(4), 313–327. <https://doi.org/10.1177/0016986217722614>
- Yue, J. (2006). Spatial Visualisation by Isometric Drawing. *Proceedings of the 2006 IJME - INTERTECH Conference*, 3.