IDA: International Design and Art Journal
ISSN: 2687-5373
Volume: 8, Issue: 1 / 2026

Research Article

Evaluation of environmental design students' sectioning skill with the Turkish version of the Santa Barbara Solids Test

Asst. Prof. Dr. Kenan Eren Şansal[1*], Asst. Prof. Güliz Taşdemir[2], Asst. Prof. Dr. Merve Şahin Kürşad[3]

[1]TED University, Faculty of Architecture and Design, Department of Interior Architecture and Environmental Design, Ankara, Türkiye.
eren.sansal@tedu.edu.tr

[2]TED University, Faculty of Architecture and Design, Department of Interior Architecture and Environmental Design, Ankara, Türkiye.
guliz.tasdemir@tedu.edu.tr

[3]TED University, Faculty of Education, Department of Educational Sciences, Ankara, Türkiye.
merve.kursad@tedu.edu.tr

*Corresponding Author

Received: 30.12.2024
Accepted: 14.01.2026

Citation:
Şansal, K. E., Taşdemir, G., Şahin Kürşad, M. (2026). Evaluation of environmental design students' sectioning skill with the Turkish version of the Santa Barbara Solids Test. *IDA: International Design and Art Journal, 8*(1), 21-35.

Abstract
Given the significance of sectioning skill in higher education and its malleability through training, it is imperative to measure this spatial skill and track its development by using an appropriate instrument. The *Santa Barbara Solids Test* is a spatial test that can be used for this purpose. However, it has been tested or used mainly with English-speaking science students. To determine whether a Turkish version of this test can be used in a culturally distinct population of design students, a two-phase study was therefore undertaken. The test's validity and reliability were initially evaluated in 173 students. Next, another sample of 200 students took the test to assess whether test scores varied by student characteristics, as in the previous studies. The results demonstrated that the psychometric properties of the Turkish version were satisfactory. A significant correlation between students' test scores and perceived spatial ability levels was found. Additionally, the test scores varied significantly by students' experience, department, and gender. From these findings, it can be concluded that the Turkish version of the test is psychometrically sound and can be used to measure and monitor the development of sectioning skill in Turkish design students.

Keywords: Design education, Environmental design, Sectioning skill, Spatial ability, The SBST

Extended Abstract

Introduction: Spatial ability is an amalgam of somewhat separate abilities or skills and plays a critical role in many professions and scientific disciplines. In addition to the importance of spatial ability, it has been demonstrated that spatial thinking and reasoning can be improved across different student populations through various training procedures. Therefore, it is essential to identify less able students using appropriate tools and to provide opportunities for them to develop their spatial skills. Given that all spatial tests may not be equally effective measures of the spatial skills critical to different fields, it is also crucial that the skills required for a specific field or domain should be measured using a relevant *domain-specific* test. One spatial skill that is more closely associated with fields such as architecture, geology, medicine, and engineering is sectioning. The *Santa Barbara Solids Test* is one of the spatial tests that can be used for measuring this skill. While this test is valid and reliable, it has been tested or used mainly with English-speaking science students. It has not been psychometrically evaluated in architecture and other environmental design students, for whom sectioning skill is deemed necessary. Moreover, it is not clear whether this test can be effectively used in culturally distinct populations.

Purpose and scope: To determine the appropriateness of this test for measuring sectioning skill in Turkish environmental design undergraduates, a study was conducted at three departments of an architecture and design faculty in Ankara, Türkiye. Since this test had not been thoroughly adapted for use in Türkiye, the test was initially translated into Turkish. Subsequently, the translated version was evaluated for validity and reliability. Moreover, student test scores were analyzed to determine whether they varied significantly by student characteristics, as in the earlier studies.

Method: The study was undertaken in two phases. In the first phase, the Santa Barbara Solids Test was translated from English to Turkish and then back to English to produce a translated version of the test. Following this translation process, 10 first-year interior design students were asked to rate whether the instructions and items of the translated version were clear. For further assessing the conceptual equivalence of the translated test, a panel of eight experts was also asked to rate the instructions and items. After making the necessary changes, a revised version of the translated test was obtained. To evaluate the content validity of individual items, the panel was asked to rate the items again. After concluding that there was no need to make additional revisions on the test, this version of the test and the *object manipulation spatial ability* subscale of the *Spatial Ability Self-Report Scale* were administered to a sample of 163 undergraduate environmental design students to assess the *construct validity, criterion validity, internal consistency,* and *convergent validity* of the obtained version. The construct validity was assessed by conducting a confirmatory factor analysis. To determine the criterion validity, the correlation between the test and the object manipulation spatial ability subscale scores was calculated. To evaluate the internal consistency and convergent validity, alpha coefficients and values for construct reliability and average variance extracted were determined. In the second phase, the test was administered to another sample of 200 undergraduate environmental design students to assess whether their scores would vary by student characteristics, namely students' year of education, department, and gender, and to conclude whether the test can be used for measuring and monitoring the development of sectioning skill in Turkish design students. To analyze the obtained test scores, one-way ANOVA for independent samples was used.

Findings and conclusion: The results of the confirmatory factor analysis that was undertaken to validate both the two- and three-factor model of the test verified the construct validity of the translated version. In addition, the significant positive correlations between participants' test scores and their reported spatial ability scores were consistent with previous research on spatial self-efficacy. They both confirmed the test's criterion validity and provided additional support for it. Moreover, the obtained Cronbach's alpha, construct reliability, and average variance extracted results were within the recommended ranges. These findings suggested that the psychometric properties of the translated version were satisfactory and that the whole test and its subscales could reliably assess sectioning skill in design students. The findings of the second phase demonstrated that the test was challenging for environmental design students. On average, the participants answered 68% of the test items correctly. Furthermore, the findings indicated that some students were significantly less successful in mentally visualizing sections. It was observed that the test scores of less- and more-experienced students differed significantly. In addition to students' experience, the test scores also varied by gender and department. While the male students scored consistently higher than their female peers, the differences in the scores reached statistical significance only for the *simple* and *embedded* subscales of the test. According to the results, the architecture students significantly outperformed their peers in the interior design and city planning departments. It was also observed that the test scores of interior design students were significantly higher than those of city planning students. Even though further research is needed to confirm these findings, several conclusions can be drawn. Firstly, it can be concluded that the translated version of the Santa Barbara Solids Test is a valid and reliable test for assessing sectioning skill in Turkish environmental design students. Secondly, some of these students may be less able to section three-dimensional objects in their minds and may be at greater risk of academic underachievement. Therefore, assessing environmental design students' sectioning skill and monitoring their progress by using the translated version would be beneficial. Thirdly, educational strategies for improving students' spatial skills by increasing their spatial self-efficacy can use this test version for sectioning skill assessment and, thus, evaluate the effectiveness of the adopted strategies. Lastly, sectioning skill can be developed through adequate departmental training for environmental design students.

Keywords: Design education, Environmental design, Sectioning skill, Spatial ability, The SBST

INTRODUCTION

A broad definition of spatial ability is "the ability to generate, retain, retrieve and transform well-structured visual images" (Lohman, 1994: 1000). A vast body of empirical evidence shows that spatial ability differs from general intelligence and that it is not a unitary construct, but rather an amalgam of somewhat separate abilities or skills (for a review, see Hegarty and Waller, 2005). While there is not a complete agreement on the structure of spatial ability (Halpern, 2012; Hegarty & Waller, 2005; Newcombe & Shipley, 2014), it is acknowledged that spatial thinking and reasoning are crucial to many professions and scientific disciplines, such as architecture, engineering, medicine, physics, and chemistry (Hegarty & Waller, 2005; Kerkman et al., 2000; Shea et al., 2001). Given its importance, it seems reasonable to assume that spatial ability is directly linked to major choice and subsequent career success. To date, longitudinal research on intellectually precocious youth has provided clear scientific evidence to support this connection. A series of studies has shown that proficiency in spatial abilities is linked to interest in science and mathematics and to pursuing a career in science and engineering (Humphreys et al., 1993; Shea et al., 2001; Webb et al., 2007). Moreover, another group of studies focusing on educational and vocational outcomes showed that earning an advanced degree in "science, technology, engineering, and mathematics" (STEM) education (Wai et al., 2009), developing innovative knowledge (Kell et al., 2013) and achieving extrinsic career success pertaining to pay and income (Lang & Kell, 2020) are related to spatial ability.

Apart from the importance of spatial ability, it has been shown that spatial thinking and reasoning can be improved across different populations through various training procedures (Uttal et al., 2013). Therefore, particular emphasis should be placed on identifying less able students and developing curricula to improve the spatial skills of those at risk. To achieve these goals, it is imperative to assess and track students' spatial abilities quantitatively using standardized tests. Given this requirement, the question arises as to whether existing spatial tests are equally effective measures of the spatial skills critical to different fields. A plausible answer to this question is that the skills needed for a specific field or domain should be measured using a relevant "domain-specific" test (Berkowitz et al., 2021; Cho & Suh, 2021; Cohen & Hegarty, 2012).

One spatial skill that is more closely associated with various STEM fields is inferring the two-dimensional (2D) sections of a three-dimensional (3D) solid object. In medicine, this skill is essential for educational attainment and for the comprehension of 2D medical images, such as X-rays (Hegarty et al., 2007; Hegarty et al., 2009). The ability to mentally visualize sections is also vital in geology, where it is termed "visual penetration ability" (Atit et al., 2015; Kali & Orion, 1996). Additionally, having trouble with sectioning 3D objects in mind is disadvantageous for engineers and is associated with poor academic performance (Cohen & Bairaktarova, 2018; Ha & Brown, 2017). Moreover, there is evidence that sectioning skill is important in architecture and more relevant to this field than other spatial skills (Berkowitz et al., 2021; Gerber et al., 2019). Given the significance of sectioning skill in several fields, it is crucial to assess it by using a suitable spatial test.

The *Santa Barbara Solids Test* (Cohen & Hegarty, 2007, 2012) is one of the spatial tests that can be used to measure sectioning skill. Even though this test has been shown to be valid and reliable, it has been tested or used mainly amongst undergraduate science, engineering, and psychology students (Cohen & Bairaktarova, 2018; Cohen & Hegarty, 2007, 2012; Ha & Brown, 2017). It has not been psychometrically evaluated in architecture and other environmental design students, for whom sectioning skill is necessary, to the best of our knowledge. Moreover, given that it is not appropriate to merely translate a measure and make it possible to use it in a different country, or more specifically, cultural context (for further information, see Beaton et al., 2000), it is not clear whether this test can be used in a culturally distinct population. Research in Western industrialized countries consistently demonstrated that males outperform females on standardized sectioning skill tests, including the Santa Barbara Solids Test (Berkowitz et al., 2021; Cohen & Hegarty, 2012; Tsutsumi et al., 2005). In these countries, it was also observed that sectioning skill varies considerably amongst university students and tends to be greater in more experienced students (Berkowitz et al., 2021; Gerber et al., 2019; Tsutsumi et al., 2005). Given these findings, it is necessary to address whether the Santa Barbara Solids Test can be effectively used to quantify possible differences in sectioning skill in another population, such as Turkish design undergraduates, and to take reasonable precautions to ensure the quality of education.

Acknowledging the importance of measuring sectioning skill and the need to determine the appropriateness of the Santa Barbara Solids Test for design students in other cultures, a study was undertaken across three departments of an architecture and design faculty in Ankara, Türkiye. Since this test had not been thoroughly adapted for use in Türkiye, the test was initially translated into Turkish. Subsequently, the translated version was evaluated for validity and reliability. Moreover, student test scores were analyzed to determine whether they varied significantly across different student characteristics. The current paper presents and discusses the results of these analyses.

METHOD

Participants

In total, 373 undergraduate students (mean age ± SD: 21.05 ± 1.74 years), studying architecture, city planning, and interior design at TED University in Ankara, Türkiye, voluntarily participated. As per the ethical clearance issued by TED University on 08 December 2023 (document number: 2023-20), all students consented to participate after being informed about the study and its objectives. Apart from speaking Turkish as their first language, there were no other inclusion criteria. At an early stage of the study, 10 first-year interior design students (5 men and 5 women) evaluated the first Turkish version of the Santa Barbara Solids Test for clarity. To assess the psychometric properties of its final Turkish version, 163 architecture and interior design students who did not participate in the first stage took the test. The descriptive statistics of these students (second-stage participants) are presented in Table 1. In the last stage, 200 students (see also Table 1 for more information on these third-stage participants) who did not participate in the previous two stages were used to evaluate the test's use among architecture and design students.

Table 1. Descriptive statistics of the second- and third-stage participants

		Second-stage participants		Third-stage participants	
		n	%	*n*	%
Gender	Female	116	71.17	145	72.50
	Male	47	28.83	55	27.50
Department	Architecture	90	54.21	84	42.00
	City planning	0	0.00	52	26.00
	Interior design	73	44.79	64	32.00
Year of education	1st	41	25.15	64	32.00
	2nd	43	26.38	59	29.50
	3rd	41	25.15	38	19.00
	4th	38	23.32	41	20.50
Total		163	100.00	200	100.00

Measures

The Santa Barbara Solids Test (SBST): A Turkish version of the SBST (Cohen & Hegarty, 2007, 2012) was used to evaluate participants' sectioning skill. The SBST is an achievement test that consists of 30 items with four answer options. The test items vary in difficulty based on the complexity of the sectioned solid and the orientation of the cutting plane. In the items, there are 10 *simple, joined,* and *embedded* solids. The *simple* solids are cylinders, cones, cubes, pyramids or prisms. The *joined* solids are made up of two simple solids that touch each other externally. *Embedded* solids are more complex solids formed by inserting two simple solids horizontally or vertically. In the test, half of the items have *orthogonal* cutting planes that are oriented perpendicularly to the vertical or horizontal axis of the solids. The remaining 15 items have *oblique* cutting planes. A total SBST score is calculated by summing the scores for all items. Alternatively, individual scores can be calculated for each sub-dimension: *simple, joined, embedded, orthogonal,* and *oblique*. In a sample of 59 psychology students, Cohen and Hegarty (2007) demonstrated that the SBST had good criterion validity and internal consistency reliability. Participants' total SBST and sub-dimension scores significantly correlated with their composite test score, calculated from two spatial ability tests. Additionally, a *Cronbach's alpha* of 0.86 was found for the overall test. In a later study with a larger sample composed mainly of undergraduate science students, Cohen and Hegarty (2012) obtained an alpha coefficient of 0.91 for the entire test and values ranging from 0.73 to 0.85 for the sub-dimensions. Moreover, participants' Scholastic Aptitude Test (SAT) mathematics scores predicted their SBST scores better than their SAT Reading scores. In addition, males significantly outperformed females in the SBST.

The Spatial Ability Self-Report Scale (SASRS): The SASRS (Turgut, 2015) was used to demonstrate convergent validity by examining correlations between participants' SBST scores and their self-reported level of *object manipulation spatial ability* (OMSA). The SASRS consists of 18 items that are rated on a 5-point scale. The responses to these items are used to quantify OMSA, *spatial navigational ability* (SNA), and *visual memory* (VM). In a sample of Turkish undergraduates, Turgut (2015) demonstrated that the SASRS had good psychometric properties. The validity analyses confirmed the scale's three-factor structure and criterion validity. Alpha coefficients for the SASRS and its OMSA, SNA, and VM subscales were reported to be 0.88, 0.88, 0.80, and 0.62, respectively. Since OMSA encompasses a wide range of skills related to spatial ability (Turgut, 2015) and can be linked with sectioning skill, only the OMSA scores were used in our analyses.

Procedure

Since there was a need to adapt the SBST for use in Türkiye, the test was adapted to the local culture and context, following the guidelines of Hambleton and Patsula (1999). Initially, two independent translators fluent in both English and Turkish translated the SBST from English into Turkish. While one translator was familiar with technical drawing terminology and was aware of the construct measured by the test, the other translator had no prior knowledge of the construct or subject matter. Both translations were reviewed by another bilingual translator and compared with the original SBST. Several inconsistencies were resolved by the translators with input from the principal investigator. After reaching a consensus, the initial translated version of the test (I-TV) was obtained. The I-TV was then translated back into English by two other translators with similar professional backgrounds. To avoid bias, both translators were completely blind to the original version of the test. Following the back-translation process, all translators and the principal investigator reviewed all translated versions of the SBST and compared them with the original version. After reconciling any discrepancies, the second translated version (S-TV) was synthesized. A sample of 10 first-year interior design students, whose first language is Turkish, was asked to rate the clarity of the S-TV's instructions and items. The instructions and items were found to be clear by 100% and 90% of the students, respectively. To further evaluate the conceptual equivalence of the S-TV, a panel of eight specialists in architectural education and one specialist in educational sciences, all native Turkish speakers, was asked to rate both the instructions and the items of the test. Since more than 80% of the panel rated the instructions as unclear, the S-TV was revised in accordance with the panel's recommendations. It was suggested to clarify that the correct answer would only represent the surface or surfaces cut by the cutting plane, as the visible surfaces beyond the cutting plane were not included in some of the answer sets. After making the necessary changes, the panel evaluated the revised S-TV (R-TV). Conceptual equivalence was achieved at this stage.

To evaluate the *content validity* of the individual items, the panel was then asked to rate the R-TV items. Based on the ratings, the items were classified as either relevant or irrelevant for computing the *item-level content validity index* (I-CVI). While the I-CVI for 29 items was 1.00, it was 0.89 for the 19th item. Furthermore, the average of the I-CVIs (S-CVI/Ave) for the items was calculated and shown to be 0.996. Given that an I-CVI of 0.78 or above for nine experts and an S-CVI/Ave of 0.90 or above are considered as satisfactory indices (Polit & Beck, 2006), it was concluded that the items did not need another revision. To further establish the psychometric properties of the R-TV, the R-TV and the OMSA subscale of the SASRS were administered to a sample of 163 students in the second stage of the study to assess *construct validity, criterion validity, internal consistency, and convergent validity.* Construct validity was assessed through a *confirmatory factor analysis.* To determine the criterion validity, the correlation between the test and OMSA subscale scores was calculated. For evaluating the internal consistency and convergent validity, alpha coefficients and values for construct reliability and average variance extracted were determined. After conducting analyses for the reliability and validity of the test, the R-TV was also administered to 200 students to demonstrate whether the test scores would vary by student characteristics in the final stage and conclude whether the test can be effectively used for measuring and monitoring the development of sectioning skill in Turkish design students. To analyze students' test scores, *one-way ANOVA for independent samples* was used. LISREL (version 8.72) was used for the confirmatory factor analysis. The remaining statistical analyses were performed by using the Statistical Package for the Social Sciences (SPSS, version 22.0).

FINDINGS

Confirmatory factor analysis (CFA)

In addition to evaluating content validity in the initial stage of the study, construct validity was assessed through a CFA. Before conducting the CFA, the suitability of the collected data for further psychometric testing was determined. First, the *Kaiser-Meyer-Olkin* (KMO) test was used to ascertain whether the sample size was adequate. There was no missing data. The KMO value for the whole sample was 0.85. Given that KMO values between 0.80 and 1.00 indicate adequacy (Shrestha, 2021), the sample size was deemed sufficient. Secondly, to identify outliers in the dataset, the raw scores for each variable were converted to *z-scores*. Since z-scores out with −3.00 or +3.00 are regarded as unsatisfactory (Yan & Su, 2009), there were no extreme values. Moreover, kurtosis and skewness values were calculated to check normality. The obtained kurtosis and skewness values ranged from −0.857 to 0.520 and from −0.587 to −0.064, respectively. Because both kurtosis and skewness values fall within the acceptable range of -2.00 to +2.00 (George & Mallery, 2020), the assumption of normality was confirmed. Since the items of the SBST were developed and categorized according to both the complexity of the sectioned solid and the cutting plane orientation, two different CFAs were performed. The first analysis was undertaken to validate the two-factor model for the two different cutting plane orientations. The CFA path diagram for the two-factor model is presented in Figure 1. Table 2 shows that the computed factor loadings for the two-dimensional structure of the adapted SBST range from 0.29 to 0.90, and that the observed variables had significant factor loadings at the 0.05 significance level in explaining the latent variables. Only Item 5 (T5) has a relatively modest loading, just below the minimum recommended loading of 0.30 (Field, 2013). Given that removing this item had a minimal effect on the overall model and that it had a significant factor loading, this item was retained.

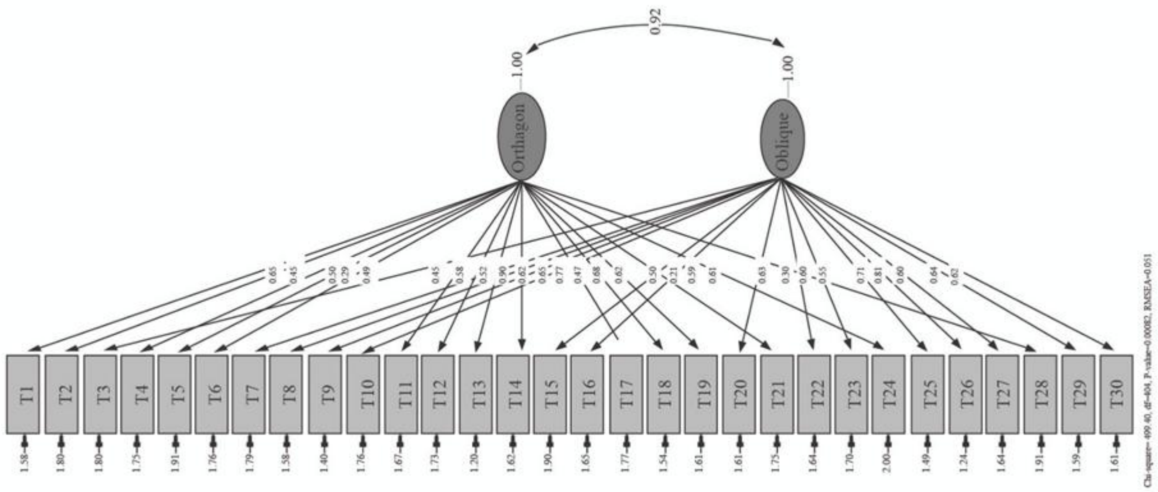

Figure 1. The obtained path diagram for the two-dimensional structure of the Turkish SBST

Table 2. Factor loadings for the two-dimensional structure of the Turkish SBST

Orthogonal		Oblique	
T1	0.65*	T3	0.45*
T2	0.45*	T7	0.46*
T4	0.50*	T8	0.65*
T5	0.29*	T9	0.77*
T6	0.49*	T10	0.49*
T11	0.58*	T15	0.31*
T12	0.52*	T16	0.59*
T13	0.90*	T20	0.63*
T14	0.62*	T22	0.60*
T17	0.47*	T23	0.55*
T18	0.68*	T25	0.71*
T19	0.62*	T26	0.81*
T21	0.50*	T27	0.60*
T24	0.61*	T29	0.64*
T28	0.30*	T30	0.63*

To assess the fit of the two-factor model, the following *goodness-of-fit* measures were employed: *chi-square to degrees of freedom ratio; comparative fit index (CFI); goodness-of-fit index (GFI); standardized root mean square residual (SRMR);* and *root mean square error of approximation.* For these measures, the recommended cut-off criteria are: $1 < x^2/df < 3$ (Carmines & McIvar, 1983); CFI ≥ 0.90 (Lai & Green, 2016); GFI > 0.90 (Schumacker & Lomax, 2010); SRMR < 0.10 (Schermelleh-Eugel et al., 2003); and RMSEA < 0.06 (Hu &

Bentler, 1999). The results of our analysis demonstrated an acceptable fit and, thus, verified the construct validity of the two-dimensional structure of the adapted SBST. All values obtained for the goodness-of-fit measures (Table 3) met the cut-off criteria.

Table 3. Goodness-of-fit analysis results for the two-factor model

X^2	df	x^2/df	CFI	GFI	SRMR	RMSEA
499.40*	404	1.23	0.90	0.94	0.07	0.05

* $p < 0.001$

A second CFA was also conducted to validate the three-factor model for the three different types of sectioned solids. The CFA path diagram is presented in Figure 2. Table 4 shows that the computed factor loadings for the two-dimensional structure of the adapted SBST range from 0.29 to 0.86, and that the observed variables had significant factor loadings at the 0.05 significance level in explaining the latent variables. Items 5 (T5) and 28 (T28) have slightly low loadings of 0.29, which are just below the minimum recommended loading. Given that removing these two items had a minimal effect on the overall model and that they had significant factor loadings, it was decided to retain them rather than modify the test.

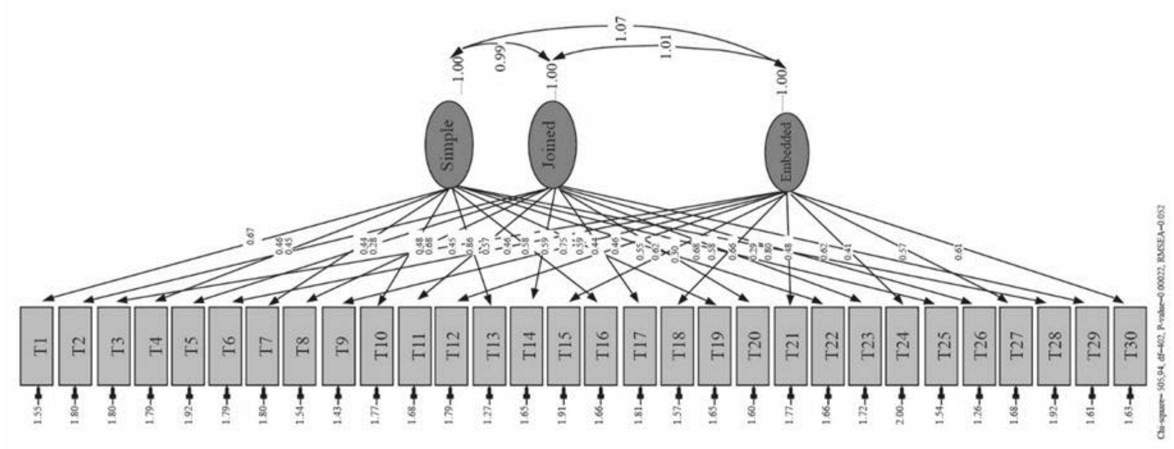

Figure 2. The obtained path diagram for the two-dimensional structure of the Turkish SBST

Table 4. Factor loadings for the three-dimensional structure of the Turkish SBST

Simple		Joined		Embedded	
T1	0.67*	T2	0.45*	T3	0.45*
T4	0.46*	T5	0.29*	T6	0.46*
T7	0.44*	T8	0.68*	T9	0.75*
T10	0.48*	T11	0.57*	T12	0.46*
T13	0.86*	T14	0.59*	T15	0.30*
T16	0.58*	T17	0.44*	T18	0.66*
T19	0.59*	T20	0.63*	T21	0.48*
T22	0.59*	T23	0.53*	T24	0.41*
T25	0.68*	T26	0.80*	T27	0.57*
T28	0.29*	T29	0.62*	T30	0.61*

* $p < 0.001$

To assess the fit of the three-factor model, the aforementioned *goodness-of-fit* measures and their cut-off criteria were used. All fit indices for the three-factor model were within the recommended ranges (Table 5). The obtained x^2/df ratio ranged from 1 to 3. While the CFI value was not lower than 0.90, the GFI value exceeded the acceptable threshold of 0.90. The SRMR and RMSEA values were both lower than the suggested cut-off values of 0.10 and 0.06, respectively. These results showed adequate fit between the data and the three-factor model tested, and they further supported the construct validity of the translated SBST.

Table 5. Goodness-of-fit analysis results for the three-factor model

X^2	df	x^2/df	CFI	GFI	SRMR	RMSEA
505.94*	402	1.25	0.90	0.95	0.07	0.05

* $p < 0.001$

Criterion validity

To determine the *criterion validity* of the adapted SBST, the correlation of participants' SBST scores with their OMSA scores was analyzed by calculating *Pearson's product-moment correlation coefficients*. Although the results of this analysis demonstrated that the obtained correlation coefficients, between 0.261 and 0.367 (Table 6), indicated either weak or moderate associations (Pallant, 2007), participants' SBST scores positively correlated with their reported spatial ability scores at statistically significant levels. These results confirmed the criterion validity of the R-TV and further supported the test's validity.

Table 6. Pearson's product-moment correlation coefficients for the relationship between the second-stage participants' (n = 163) SBST and OMSA scores

	SBST total	SBST simple	SBST joined	SBST embedded	SBST orthogonal	SBST oblique
OMSA	0.357*	0.261*	0.325*	0.367*	0.283*	0.350*

* $p < 0.05$

Internal consistency and convergent validity

While the *Kuder-Richardson 20* (KR-20) is generally used for estimating the reliability of dichotomous measures (Villaume & Weaver, 1996), Cronbach's alpha can also be used for binary items and yields the same result as the KR-20 (Meyer, 2010). Given the suitability of utilizing "Cronbach's alpha" for the adapted SBST, alpha coefficients were calculated for the whole test and its subscales. It can be seen from Table 7 that all alpha coefficients for the adapted test are not below the threshold value of 0.70 (Taber, 2018) and, thus, indicate the internal consistency of the adapted SBST. It is also evident from Table 7 that the alpha coefficients for the adapted test are comparable to those obtained for the original test, except for the coefficient for the *orthogonal* subscale.

Table 7. "Cronbach's alpha" coefficients, CRs, and AVEs for the SBST and its subscales

	SBST total	SBST simple	SBST joined	SBST embedded	SBST orthogonal	SBST oblique
alpha	0.85	0.76	0.78	0.70	0.70	0.81

	Two-dimensional structure			Three-dimensional structure			
	SBST orthogonal	SBST oblique	SBST total	SBST simple	SBST joined	SBST embedded	SBST total
CR	0.97	0.98	0.99	0.92	0.92	0.89	0.91
AVE	0.31	0.36	0.34	0.34	0.33	0.28	0.32

Since Cronbach's alpha may underestimate reliability (Hair et al., 2019), the construct reliability (CR) value, considered to be a more suitable alternative (Hair et al., 2019; Netemeyer et al., 2003), was also calculated. In addition to CR value, *average variance extracted* (AVE), a better criterion for internal consistency or stability (Netemeyer et al., 2003), was computed as well. Table 7 also shows that all CR values range from 0.89 to 0.99. Given that these values are above the acceptable CR threshold of 0.70 (Hair et al., 2019), they indicate an adequate convergence or internal consistency. Table 7 also shows that the AVE values range from 0.28 to 0.36. Although the minimum value for AVE is 0.50 (Hair et al., 2019), these values are still acceptable for two reasons. Firstly, for a new measure, an AVE of 0.25 or higher can be considered satisfactory (Wang et al., 2023). Secondly, if CR values exceed the threshold, an AVE below 0.50 is still deemed adequate and an indicator of convergent validity (Shrestha, 2021). Therefore, it is reasonable to conclude that the Turkish version of the SBST has an acceptable internal consistency and convergent validity.

To further confirm our findings on the internal consistency of the adapted SBST, alpha coefficients for the whole test and its subscales were calculated from the dataset for the 200 third-stage participants. Before this analysis, the adequacy of sampling was determined by the KMO test. There was no missing data. A KMO value of 0.85 was found for the whole sample. Since this value fell within the range of 0.80 to 1.00, the sample size was deemed adequate. Moreover, to identify outliers in the dataset, the raw scores for each variable were converted to *z-scores*. Given that the obtained z-scores were between −3.00 and +3.00, there were no extreme values. Also, kurtosis and skewness values were calculated to check normality. The obtained kurtosis and skewness values ranged from −0.876 to 0.342 and from −1.506 to 0.172, respectively. Because both kurtosis and skewness values were within the range of −2.00 to +2.00, the assumption of normality was confirmed. The reliability analysis demonstrated that all alpha coefficients were above the 0.70 threshold and thus indicated satisfactory reliability. The coefficient for the whole scale was 0.87. The coefficients for the *simple, joined, embedded, orthogonal,* and *oblique* subscales were 0.75, 0.71, 0.72, 0.79, and 0.81, respectively.

Test score differences in the participants

To assess the use of the R-TV among Turkish design students, we also examined whether third-stage participants' test scores varied significantly by year of education, department, and gender. The descriptive statistics for participants' correct answers in the R-TV are given in Table 8. On average, the participants answered 68% of the R-TV items correctly. While they successfully answered 79% of the items with *orthogonal* cutting planes, they correctly answered 56% of the items with *oblique* cutting planes. The success rate for the items with *simple, joined,* and *embedded* solids were 68%, 70%, and 65%, respectively. Approximately 83% (165) of students had an average score of 15 or more.

Table 8. Descriptive statistics for the correct answers in the R-TV

	Mean	*SD*	%
SBST simple	6.76	2.48	67.60
SBST joined	6.97	2.31	69.70
SBST embedded	6.48	2.14	64.80
SBST orthogonal	11.86	3.03	79.07
SBST oblique	8.35	3.83	55.67
SBST total	20.51	6.23	68.37

To ascertain whether the SBST scores of less and more experienced students significantly differed, participants' test scores were analyzed by using the "one-way ANOVA for independent samples". The analysis results are presented in Table 9. The results indicated that the scores for the whole test [$F(3,196) = 9.780$, $p < 0.001$] and *simple* [$F(3,196) = 5.995$, $p < 0.001$], *joined* [$F(3,196) = 9.150$, $p < 0.001$], *embedded* [$F(3,196) = 8.362$, $p < 0.001$], *orthogonal* [$F(3,196) = 8.446$, $p < 0.001$] and *oblique* [$F(3,196) = 7.865$, $p < 0.001$] subscales varied significantly by the year of education or students' experience. According to the post-hoc test results (Table 9), the more experienced students significantly outperformed their less experienced peers. The fourth-year students had the highest test scores. They were followed by the third-, second-, and first-year students, respectively.

Table 9. One-way ANOVA and post-hoc test results for students' year of education

		Sum of squares	*df*	Mean square	*F*	Significant differences
SBST simple	Between groups	103.07	3	34.35	5.995*	$1^{st} < 2^{nd} < 3^{rd} < 4^{th}$ year
	Within groups	1123.40	196	5.73		
	Total	1226.48	199			
SBST joined	Between groups	131.55	3	43.85	9.150*	$1^{st} < 2^{nd} < 3^{rd} < 4^{th}$ year
	Within groups	939.32	196	4.79		
	Total	1070.87	199			
SBST embedded	Between groups	104.15	3	34.71	8.362*	$1^{st} < 2^{nd} < 3^{rd} < 4^{th}$ year
	Within groups	813.76	196	4.15		
	Total	917.92	199			
SBST orthogonal	Between groups	209.64	3	69.88	8.446*	$1^{st} < 2^{nd} < 3^{rd} < 4^{th}$ year
	Within groups	1621.71	196	8.27		
	Total	1831.35	199			
SBST oblique	Between groups	314.98	3	104.99	7.865*	$1^{st} < 2^{nd} < 3^{rd} < 4^{th}$ year
	Within groups	2616.51	196	13.35		
	Total	2931.50	199			
SBST total	Between groups	1007.20	3	335.73	9.780*	$1^{st} < 2^{nd} < 3^{rd} < 4^{th}$ year
	Within groups	6728.55	196	34.32		
	Total	7735.75	199			

* $p < 0.001$

The obtained scores on the R-TV and its subscales were also analyzed by using the one-way ANOVA for independent samples to assess the possible effects of different design fields on participants' sectioning skill. The analysis results are presented in Table 10. The scores for the whole test [$F(2,197) = 14.280$, $p < 0.001$] and *simple* [$F(2,197) = 12.297$, $p < 0.001$], *joined* [$F(2,197) = 9.473$, $p < 0.001$], *embedded* [$F(2,197) = 12.493$, $p < 0.001$], *orthogonal* [$F(2,197) = 8.778$, $p < 0.001$] and *oblique* [$F(2,197) = 14.325$, $p < 0.001$] subscales varied significantly by participants' department. According to the post-hoc test results (Table 10), the architecture students significantly outperformed the interior design and city planning students. Also, the SBST scores of the interior design students were significantly better than those of the city planning students.

Table 10. One-way ANOVA and post-hoc test results for the students' department

		Sum of squares	*df*	Mean square	*F*	Significant differences
SBST simple	Between groups	136.12	2	68.06	12.297*	City planning < Interior design < Architecture
	Within groups	1090.36	197	5.53		
	Total	1226.48	199			
SBST joined	Between groups	93.95	2	46.97	9.473*	City planning < Interior design < Architecture
	Within groups	976.92	197	4.95		
	Total	1070.875	199			
SBST embedded	Between groups	103.31	2	51.65	12.493*	City planning < Interior design < Architecture
	Within groups	814.60	197	4.13		
	Total	917.92	199			
SBST orthogonal	Between groups	150.01	2	75.00	8.788*	City planning < Interior design < Architecture
	Within groups	1681.34	197	8.53		
	Total	1831.35	199			
SBST oblique	Between groups	372.20	2	186.10	14.325*	City planning < Interior design < Architecture
	Within groups	2559.29	197	12.99		
	Total	2931.50	199			
SBST total	Between groups	979.51	2	489.75	14.280*	City planning < Interior design < Architecture
	Within groups	6756.24	197	34.29		
	Total	7735.75	199			

* $p < 0.001$

To determine whether participants' SBST scores significantly differed between the male and female participants, the test scores were analyzed by using the independent samples t-test. The analysis results are presented in Table 11. While the males scored consistently higher than their female peers, the differences in the scores reached statistical significance only for the *simple* [t(198) = −2.071, p = 0.040] and *embedded* [t(198) = −2.127, p = 0.035] subscales.

Table 11. Independent samples t-test results for students' gender

		n	Mean	*SD*	Mean difference	*t*	*df*	*p*
SBST simple	Female	145	6.53	2.42	−0.80	−2.071	198	0.040
	Male	55	7.34	2.54				
SBST joined	Female	145	6.91	2.39	−0.21	−0.571	198	0.569
	Male	55	7.12	2.11				
SBST embedded	Female	145	6.28	2.05	−0.71	−2.127	198	0.035
	Male	55	7.00	2.30				
SBST orthogonal	Female	145	11.71	3.10	−0.56	−1.172	198	0.243
	Male	55	12.27	2.82				
SBST oblique	Female	145	8.02	3.64	−1.17	−1.942	198	0.054
	Male	55	9.20	4.22				
SBST total	Female	145	19.73	6.12	−1.73	−1.766	198	0.079
	Male	55	21.47	6.40				

CONCLUSION

This study highlights the relevance of spatial ability and skills for academic achievement and learning. The objectives of this study were both exploratory and confirmatory. Since a psychometrically sound Turkish version of the SBST was unavailable, the first objective was to translate it. The second objective was to evaluate the validity and reliability of the Turkish version of the test in a sample of Turkish undergraduates. Given the absence of evidence on the use of the SBST among environmental design students for whom sectioning skill is crucial, a sample of architecture, city planning, and interior design students was utilized. In addition to assessing the adapted test for validity and reliability, another objective was to determine whether specific student characteristics could explain any marked differences in SBST scores.

The analyses verified the validity and reliability of the Turkish version of the SBST. The CFA results demonstrated that items with three different solid types and those with two different cutting planes loaded on separate factors. These results confirm both the three-factor and two-factor structures of the test, for which its items were developed and categorized according to the complexity of the sectioned solid and cutting-plane orientation, and thus the test's construct validity. Furthermore, students' perceptions of their spatial ability, or more specifically, spatial self-efficacy, were significantly correlated with their actual performance on the SBST. This finding aligns with previous research on spatial self-efficacy (Power et al., 2016; Safadel et al., 2023; Towle et al., 2005) and provides additional support for the validity of the adapted SBST. The reliability results suggest that the entire adapted test and its subscales can reliably assess sectioning skill in our sample. Moreover, the results demonstrated that the adapted SBST was a challenging test for the participants. On average, our third-stage participants answered roughly 20 of the 30 items correctly. This finding is comparable with that of Cohen and Hegarty (2012). They found that about 68% of the answers were correct in their sample, mainly composed of science and engineering students. Apart from the use of the adapted test in environmental design students, the results point to the importance of boosting students' spatial self-efficacy. This requires a holistic approach for improving spatial ability rather than focusing solely on students' actual ability. In addition to providing opportunities to develop students' spatial skills, incorporating educational strategies, such as offering help and giving positive feedback on students' performance, can be more beneficial and thus more appropriate.

The results also indicated that some students were significantly less successful in mentally visualizing the 2D sections of the 3D solids. The subscale and whole test scores of the students varied significantly by the year of education. This finding is consistent with those of the earlier studies on the differences in sectioning skills between experienced and inexperienced students (Berkowitz et al., 2021; Gerber et al., 2019; Tsutsumi et al., 2005) and the influence of design training and instruction on the development of spatial cognition (Berkan et al., 2020; Gomez-Tone et al., 2021; Lin, 2016). In addition to student experience, the SBST scores also varied by gender. Although all SBST scores for male participants were higher than those of their female peers, the significant differences were limited to the "simple" and "embedded" subscales. These findings offer some empirical support for the previously reported outperformance of males relative to females in sectioning skill tests (Berkowitz et al., 2021; Cohen & Hegarty, 2012; Tsutsumi et al., 2005), and they align with Cohen and Hegarty's (2012) findings. While the root cause of these differences remains to be elucidated and warrant further investigation, it has been proposed that the combined effects of various biological, psychological and societal factors can account for the observed differences in spatial skills and other cognitive abilities (for this discussion, see Halpern, 2012) and that "the co-construction of gender and spatial ability" can also explain the gender differences (Bartlett & Camba, 2023). It is important to note that we should not focus solely on these differences but also on a much more fundamental issue. Given that both genders can benefit from training and develop their spatial skills equally (Baenninger & Newcombe, 1989; Newcombe et al., 2002; Uttal et al., 2013), future studies should place a greater emphasis on how to maximize spatial competence in both genders to facilitate learning and enhance academic performance. Another finding of the present study that warrants further exploration is the observed differences in SBST performance among students from the three departments. While it needs to be confirmed, the significant differences between the city planning students and those of architecture and interior design students may stem from the curricula of the degree programs. In addition to the compulsory technical drawing and design courses available across all programs, a wide variety of both compulsory and elective courses in architectural design-related subjects, such as structure and construction technologies, are offered to architecture and interior design students. Since these extra courses may have fostered spatial ability development, the SBST performance of the city planning students was found to be comparatively low. Given the demonstrated relationship between mathematics achievement and overall SBST score (Cohen & Hegarty, 2012), one reason for the differences between architecture and other environmental design students may be the admission criteria used for architecture programs in Türkiye. Unlike the interior design and city planning programs, which admit students only based on their university entrance examination scores, the architecture programs also require students to rank among the top 250,000 examinees. Since almost one-third of the examination is composed entirely of mathematics questions, architecture students can be expected to be relatively more successful in mathematics and, thus, in the SBST.

Several limitations of our study should be noted here and considered in interpreting our findings and planning future research. Firstly, the study sample comprised environmental design students from a single faculty. To generalize the findings to other design- and STEM-related disciplines, for which sectioning skill is particularly important, additional research across diverse samples appears beneficial. Secondly, the sample was composed mainly of females due to the scarcity of males in those departments. While this imbalance has been reported in the psychometric assessment of the SBST amongst undergraduate science students (Cohen & Hegarty, 2012) and thus is not unexpected in design-related disciplines, possible gender effects on the study outcomes should also be considered and investigated further. Lastly, the development of sectioning skill in our sample was determined by comparing students' SBST scores across different years of study. Even though this comparison provides some support for the benefits of training in spatial skills, monitoring students' progress by measuring their sectioning skill at multiple points may further improve our understanding of training effects.

Evidently, there is insufficient empirical evidence to fully understand the impact of sectioning skill and other spatial skills on educational outcomes, or to evaluate the potential use of the SBST for quantifying sectioning skill across different student populations. Nevertheless, we can draw some conclusions from the existing findings. Firstly, it can be concluded that SBST's Turkish version is a psychometrically sound test. Secondly, some students may encounter greater difficulty visualizing 2D sections of 3D objects than their peers and be at a greater risk of academic underachievement. Therefore, it can also be concluded that these students need to be identified and monitored by using standardized instruments, such as the SBST, to take all reasonable precautions. Thirdly, future attempts to improve students' spatial self-efficacy can use this version of the SBST as a diagnostic tool to identify students who are not highly skilled at sectioning and, thus, are likely to benefit more from educational strategies to boost spatial self-efficacy. Lastly, given that students' SBST scores varied significantly by the year of education, sectioning skill can be developed through adequate departmental training in environmental design students.

Authors' Contributions

All authors contributed to the study conception and design. Material preparation and data collection were performed by Kenan Eren Şansal and Güliz Taşdemir. Data analysis was performed by Merve Şahin Kürşad. The first draft of the manuscript was written by Kenan Eren Şansal and Merve Şahin Kürşad and edited by Güliz Taşdemir. All authors commented on the previous versions of the manuscript. All authors read and approved the final manuscript.

Competing Interests

There is no potential conflict of interest.

Ethics Committee Declaration

Ethics committee approval dated 08/12/2023 and numbered 2023-20 was obtained by Human Research Ethics Committee of TED University.

REFERENCES

Atit, K., Gagnier, K., & Shipley, T. F. (2015). Student gestures aid penetrative thinking. *Journal of Geoscience Education, 63*(1), 66-72. https://doi.org/10.5408/14-008.1

Baenniger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles, 20,* 327-344. https://doi.org/10.1007/BF00287729

Bartlett, K. A., & Camba, J. D. (2023). Gender differences in spatial ability: A critical review. *Educational Psychology Review 35*(8), Article 8. https://doi.org/10.1007/s10648-023-09728-2

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186-3191.

Berkan, S. T., Oztas, S. K., Kara, F. İ., & Vardar, A. E. (2020). The role of spatial ability on architecture education. *Design and Technology Education: An International Journal, 25*(3), 103-126.

Berkowitz, M., Gerber, A., Thurn, C. M., Emo, B., Hoelscher, C., & Stern, E. (2021). Spatial abilities for architecture: Cross sectional and longitudinal assessment with novel and existing spatial ability tests. *Frontiers in Psychology*, *11*, Article 609363. https://doi.org/10.3389/fpsyg.2020.609363

Carmines, E. G., & McIver, J. P. (1983). An introduction to analysis of models with unobserved variables. *Political Methodology*, *9*(1), 51-102.

Cho, J. Y., & Suh, J. (2021). The architecture and interior design domain-specific spatial ability test (AISAT): Its validity and reliability. *Journal of Interior Design*, *47*(2), 11-30. https://doi.org/10.1111/joid.1221

Cohen, C., & Bairaktarova, D. (2018). A cognitive approach to spatial visualization assessment for first-year engineering students. *Engineering Design Graphics Journal*, *82*(3), 1-19. http://www.edgj.org/index.php/EDGJ/article/view/702

Cohen, C. A., & Hegarty, M. (2007). Sources of difficulty in imagining cross sections of 3D objects. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the Annual Meeting of Cognitive Science Society* (pp. 179-184). Cognitive Science Society.

Cohen, C. A., & Hegarty, M. (2012). Inferring cross sections of 3D objects: A new spatial thinking test. *Learning and Individual Differences*, *22*(6), 868-874. https://doi.org/10.1016/j.lindif.2012.05.007

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*. Sage.

George, D., & Mallery, P. (2020). *IBM SPSS statistics 26: Step by step: A simple guide and reference*. Routledge. https://doi.org/10.4324/9780429056765

Gerber, A., Berkowitz, M., Emo, B., Kurath, S., Holscher, C., & Stern, E. (2019). Does space matter? A cross-disciplinary investigation upon spatial abilities of architects. In C. Leopold, C. Robeller & U. Weber (Eds.), *Research Culture in Architecture: Cross-Disciplinary Collaboration* (pp. 121-169). Birkhauser.

Gomez-Tone, H. C., Martin-Gutierrez, J., Bustamante-Escapa, J., & Bustamante-Escapa, P. (2021). Spatial skills and perceptions of space: Representing 2D drawings as 3D drawings inside immersive virtual reality. *Applied Sciences*, *11*(4), Article 1475. https://doi.org/10.3390/app11041475

Ha, O., & Brown, S. (2017, June 24–28). *Spatial Reasoning Difference between Civil and Mechanical Engineering Students in Learning Mechanics of Materials Course: A Case of Cross-Sectional Inference* [Conference Presentation]. ASEE Annual Conference & Exposition, Columbus, OH, United States. https://peer.asee.org/28836

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multi Variate Data Analysis*. Cengage.

Halpern, D. F. (2012). *Sex differences in cognitive abilities*. Psychology Press. https://doi.org/10.4324/9780203816530

Hambleton, R. K., & Patsula, L. (1999). Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, *1*(1), 1-13.

Hegarty, M., Keehner, M., Cohen, C., Montello, D. R., & Lippa, Y. (2007). The role of spatial cognition in Medicine: Applications for selecting and training professionals. In G. L. Allen (Ed.), *Applied Spatial Cognition from Research to Cognitive Technology* (pp. 285-315). Psychology Press. https://doi.org/10.4324/9781003064350

Hegarty, M., Keehner, M., Khooshabeh, P., & Montello, D. R. (2009). How spatial abilities enhance and are enhanced by dental education. *Learning and Individual Differences*, *19*(1), 61-70. https://doi.org/10.1016/j.lindif.2008.04.006

Hegarty, M., & Waller, D. A. (2005). Individual differences in spatial abilities. In P. Shah & A. Miyake (Eds.), *The Cambridge Handbook of Visuospatial Thinking* (pp. 121-169). Cambridge University Press. https://doi.org/10.1017/CBO9780511610448.005

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *Journal of Applied Psychology*, *78*(2), 250-261. https://doi.org/10.1037/0021-9010.78.2.250

Kali, Y., & Orion, N. (1996). Spatial abilities of high-school students in the perception of geologic structures. *Journal of Research in Science Teaching*, *33*(4), 369-391. https://doi.org/10.1002/(SICI)1098-2736(199604)33:4<369::AID-TEA2>3.0.CO;2-Q

Kell, H. J., Lubinski, D., Benbow, C. P., & Steiger, J. H. (2013). Creativity and technical innovation: Spatial ability's unique role. *Psychological Science*, *24(*9), 1831-1836. https://doi.org/10.1177/09567976134786

Kerkman, D. D., Wise, J. C., & Harwood, E. A. (2000). Impossible "mental rotation" problems: A mismeasure of of women's spatial abilities? *Learning and Individual Differences*, *12*(3), 253-269. https://doi.org/10.1016/S1041-6080(01)00039-5

Lai, K., & Green, s. B. (2016). The problem with having two watches: Assessment of fit when RMSEA and CFI disagree. *Multivariate Behavioral Research*, *51*(2-3), 220-239. https://doi.org/10.1080/00273171.2015.1134306

Lang, J. W. B., & Kell, H. J. (2020). General mental ability and specific abilities: Their relative importance for extrinsic career success. *Journal of Applied Psychology*, *105*(9), 1047-1061. https://doi.org/10.1037/apl0000472

Lin, H. (2016). Influence of design training and spatial solution strategies on spatial ability performance. *International Journal of Technology and Design Education*, *26,* 123-131. https://doi.org/10.1007/s10798-015-9302-7

Lohman, D. F. (1994). Spatial ability. In R. J. Sternberg (Ed.), *Encyclopaedia of Human Intelligence* (pp. 1000-1007). Macmillan Publishing Company.

Meyer, J. P. (2010). *Understanding measurement: Reliability*. Oxford University Press.

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Sage. https://doi.org/10.4135/9781412985772

Newcombe, N. S., Mathason, L., & Terlecki, M. (2002). Maximization of spatial competence: More important than finding the cause of sex differences. In A. McGillicuddy-De Lisi & R. De Lisi (Eds.), *Biology, Society, and Behavior: The Development of Sex Differences in Cognition* (pp. 183-206). Ablex Publishing.

Newcombe, N. S., & Shipley, T. F. (2014). Thinking about spatial thinking: New typology, new assessments. In J. S. Gero (Ed.), *Studying Visual and Spatial Reasoning for Design Creativity* (pp. 179-192). Springer. https://doi.org/10.1007/978-94-017-9297-4_10

Pallant, J. (2007). SPSS survival manual: *A step by step guide to data analysis using SPSS for Windows*. Open University Press.

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489-497. https://doi.org/10.1002/nur.20147

Power, J., Buckley, J., & Seery, N. (2016, January 24–26). *Visualizing Success: Investigating the Relationship between Ability and Self-Efficacy in the Domain of Visual Processing* [Conference Presentation]. ASEE EDGD 70th Midyear Conference, Daytona Beach, FL, United States. https://commons.erau.edu/asee-edgd/conference70/

Safadel, P., White, D., & Kia, A. (2023). Spatial self-efficacy and spatial ability: An analysis of their relationship. *New Review of Hypermedia and Multimedia*, *29*(2), 114-150. https://doi.org/10.1080/13614568.2023.2248057

Schermelleh-Engel, K., Moosbrugger, H., & Muller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8*(2), 23-74. https://doi.org/10.23668/psycharchives.12784

Schumacker, R. E., & Lomax, R. G. (2010). *A Beginner's Guide to Structural Equation Modelling*. Routledge.

Shea, D. L., Lubinski, D., & Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: A 20-year longitudinal study. *Journal of Educational Psychology*, *93*(3), 604-614. https://doi.org/10.1037/0022-0663.93.3.604

Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American Journal of Applied Mathematics and Statistics*, *9*(1), 4-11. https://doi.org/10.12691/ajams-9-1-2

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48, 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Towle, E., Mann, J., Kinsey, B., O'Brien, E. J., Bauer, C. F., & Champoux, R. (2005, October 19–22). *Assessing the Self-Efficacy and Spatial Ability of Engineering Students from Multiple Disciplines* [Conference Presentation]. 35th ASEE/IEEE Frontiers in Education Conference, Indianapolis, IN, United States.

Tsutsumi, E., Schrocker, H. P., Stachel, H., & Weiss, G. (2005). Evaluation of students' spatial abilities in Austria and Germany. *Journal for Geometry and Graphics*, *9*(1), 107-117.

Turgut, M. (2015). Development of the Spatial Ability Self-report Scale (SASRS): Reliability and validity studies. *Quality and Quantity*, *49,* 1997-2014. https://doi.org/10.1007/s11135-014-0086-8

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

Villaume, W. A., & Weaver, J. B. (1996). A factorial approach to establish reliable listening measures from the WBLOT and the KCLT: Full information factor analysis of dichotomous data. *International Journal of Listening*, *10*(1), 1-20. https://doi.org/10.1207/s1932586xijl1001_1

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

Wang, B., Rau, P. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour and Information Technology*, *42*(9), 1324-1337. https://doi.org/10.1080/0144929X.2022.2072768

Webb, R. M., Lubinski, D., & Benbow, C. P. (2007). Spatial ability: A neglected dimension in talent searches for intellectually precocious youth. *Journal of Educational Psychology*, *99*(2), 397-420. https://doi.org/10.1037/0022-0663.99.2.397

Yan, X., & Su, X. G. (2009). *Linear regression analysis: Theory and computing*. World Scientific. https://doi.org/10.1142/6986

Authors' Biographies

Kenan Eren Şansal continued his studies at University College London (UCL) in UK after receiving his BFA in interior architecture and environmental design from ID Bilkent University in 2004. He received his MSc and PhD degrees in architectural lighting design from UCL in 2006 and 2012, respectively. He is currently working as an assistant professor at TED University and teaching various undergraduate courses on interior design. He is pursuing his research studies on architectural lighting design and interior design education.

Güliz Taşdemir completed her undergraduate degree at Başkent University, Department of Interior Architecture and Environmental Design. She received her both master's and doctorate degree from Hacettepe University. She studies alternative stories of the city through oral history, discusses identity/meaning/context relations and continues her research on the connection between space and social history. She is currently working an assistant professor at TED University and teaching various undergraduate courses on interior design. Dr. Taşdemir is also an active member of 4T Design and Design History Association and Docomomo_Turkey, Interior Design Committee.

Merve Şahin Kürşad graduated from the Department of Mathematics and Science Education at Middle East Technical University in 2011. She received her master's degree from Bolu Abant İzzet Baysal University and her doctorate degree from Ankara University in 2014 and 2020, respectively. She is currently working as an assistant professor at TED University and teaching various courses on research methods in education, test construction and 21st century skills for teachers. Her research interests mainly focus on measurement and evaluation in education.