

Diagnostic Assessment of Spatial Intelligence in Iranian English-major University Students: A Psychometric Study

¹ Saeed Esmailnia

² Noorali Farrokhi*

³ Fatemeh Shaterian

Research Paper

IJEAP-2605-2201

Received: 2026-05-15

Accepted: 2026-06-29

Published: 2021-06-30

Abstract: Spatial intelligence is critical for comprehending STEM-related English for Academic Purposes (EAP) materials, yet few diagnostic tools assess fine-grained spatial abilities among English-major students, and gender differences remain underexplored. This study developed and validated a spatial intelligence test based on cognitive diagnostic models (CDM) and examined gender differences in five components: mental rotation, spatial visualization, understanding of 3D relationships, spatial perceptual speed, and visual memory. Using an exploratory mixed-methods design, the test was initially piloted on a general high school sample ($N = 250$) to establish basic psychometric properties, then re-administered to a separate validation sample of 98 Iranian English-major university students (52 male and 46 female). Content validity ($CVR = 0.80-1.00$; $CVI = 0.80-1.00$), exploratory factor analysis (five factors explaining 59.18% of variance), confirmatory factor analysis ($\chi^2/df = 1.66$, $RMSEA = 0.047$, $CFI = 0.95$, $TLI = 0.95$), and internal consistency (Cronbach's $\alpha = 0.85-0.90$) confirmed strong psychometric properties. The G-DINA model was fitted. Gender differences significantly favored males in mental rotation (77% vs. 63%), spatial visualization (74% vs. 63%), and understanding of 3D relationships (64% vs. 51%), but not in spatial perceptual speed or visual memory. The test provides a valid, reliable, culturally adapted diagnostic tool for EAP contexts. Implications include targeted visual interventions (e.g., flowchart tasks, annotated 3D diagrams) and visual feedback reports for personalized instruction.

Keywords: Cognitive Diagnostic Models, English for Academic Purposes (EAP), English-major Students, Gender Differences, Spatial Intelligence

Introduction

In the contemporary landscape of English for Academic Purposes (EAP), success is increasingly defined not only by linguistic proficiency but also by the underlying cognitive abilities that facilitate academic processing (Hulur et al., 2016). While traditional EAP curricula prioritize lexical and grammatical competence, emerging research suggests that spatial intelligence plays a pivotal role in comprehending complex academic materials, particularly in STEM-related English contexts where visual literacy is paramount (Wu et al., 2025). Spatial intelligence affects EAP learning by enabling learners to mentally

¹Ph.D. in Assessment and Measurement, Email: saeedesmailnia@gmail.com, Department of Psychology, Faculty of Psychology and Education, Islamic Azad University, Saveh Branch, Saveh, Iran.

²Professor of Educational Psychology, Email: farrokhinoorali@gmail.com, Department of Assessment and Measurement, Faculty of Psychology and Education, Allameh Tabataba'i University, Tehran, Iran.

³Assistant Professor in Educational Psychology, Email: fshaterian@yahoo.com, Department of Educational Measurement, Faculty of Humanities, Islamic Azad University, Saveh Branch, Saveh, Iran.

manipulate and organize visually presented information, such as diagrams, flowcharts, graphs, and three-dimensional models commonly found in science and engineering textbooks. Specific EAP tasks that rely on spatial cognition include interpreting research figures, understanding spatial prepositions and locative expressions in technical descriptions, following multi-step procedural instructions accompanied by illustrations, and integrating textual information with graphical data (e.g., in reading-to-write tasks) (Nagy-Kondor, 2007; Ouyang et al., 2022). Consequently, neglecting these cognitive underpinnings in language assessment may result in an incomplete understanding of learner potential, especially in educational systems aiming to prepare students for global academic challenges (Tahir et al., 2025).

Despite the recognized importance of cognitive skills, there remains a significant gap in diagnostic tools capable of dissecting the specific attributes contributing to academic performance within English-major contexts. Classical test theories often provide a unidimensional score that obscures the nuanced strengths and weaknesses of learners, thereby limiting the effectiveness of targeted pedagogical interventions (De la Torre & Sorrel, 2023). This limitation is particularly critical when considering the diverse cognitive profiles of learners in non-Western contexts, where standardized Western instruments may not fully capture the local educational realities (Mpiti et al., 2025). To address this gap, Cognitive Diagnostic Models (CDM) have emerged at the intersection of cognitive psychology and psychometrics, designed to provide fine-grained diagnostic feedback about learners' mastery of specific subskills (De la Torre, 2009; Rupp & Templin, 2007). Unlike classical test theory, which typically yields a single total score, CDMs classify examinees into attribute mastery profiles, indicating which underlying cognitive skills an individual has acquired (e.g., mental rotation, spatial visualization) and which still require further development (De la Torre & Sorrel, 2023). Thus, CDMs offer a robust framework for identifying mastery levels across specific cognitive attributes, providing a detailed analysis that aligns with the growing demand for personalized learning pathways in language education (Rupp & Templin, 2007; Wafa et al., 2023).

Spatial intelligence is conceptualized as an independent and vital form of intelligence that defines an individual's capacity to visualize objects, comprehend spatial relationships, and mentally rotate shapes within visual environments (Gardner, 2020; Nagy-Kondor, 2007). It encompasses spatial perception and intuition, mental imagery, mental rotation of forms, understanding of spatial relations, and recognition of spatial locations (Nagy-Kondor, 2007). Spatial intelligence is also regarded as a human computational capacity that reflects the mental ability or skill required to solve spatial problems related to orientation, visualizing objects from different angles and positions, recognizing faces or scenes, and attending to fine-grained details (Nagy-Kondor & Esmailnia, 2021a, 2021b, 2022; Suprpto et al., 2018). In other words, it refers to the ability to construct and manipulate internal representations of space, imagine spatial configurations from multiple perspectives, attend to subtle features and details, and recognize visual scenes even in the absence of direct physical stimuli.

The intersection of cognitive abilities and gender dynamics presents a complex variable in educational assessment that requires careful examination through modern diagnostic lenses (Beikian & Esmailnia, 2023). Recent studies indicate that gender differences in spatial tasks are not merely biological but are significantly mediated by environmental factors, stereotypes, and educational opportunities (Ebert et al., 2025; Ghasemi et al., 2025). In the realm of English-major, where course materials and instructional strategies may inadvertently reinforce gendered cognitive roles, understanding these differences is crucial for promoting equity (Dahmardeh et al., 2025; Mahmoudikia & Ahmadi, 2020). For instance, if female learners exhibit different patterns of spatial mastery compared to their male counterparts, EAP instructors may need to consider adapting their visual teaching materials to ensure inclusive comprehension across all student demographics (Fioriti et al., 2024).

Despite the availability of international instruments for assessing spatial intelligence—such as the Purdue Spatial Visualization Test (PSVT), the Mental Cutting Test (MCT), and the Differential Aptitude Test – Space Relations (DAT-SR)—a significant research gap persists regarding their applicability for Iranian high school and university students, particularly within English language learning contexts. These

instruments, while well-validated in Western populations, have not been systematically examined for cultural or linguistic appropriateness in Iran. For example, none have been translated, adapted, or normed for Persian-speaking English-major learners, nor have they been calibrated to detect fine-grained spatial skill profiles using cognitive diagnostic models. There is, therefore, a notable scarcity of culturally validated tools that can provide diagnostic profiles of learners' cognitive strengths and weaknesses across specific spatial dimensions (e.g., mental rotation, spatial visualization, understanding of 3D relationships). Consequently, conducting research of this kind may help establish a reliable, bias-sensitive baseline of cognitive readiness that complements linguistic assessments.

The significance of this study lies not only in providing a foundation for future investigations into the relationship between spatial cognition and language learning success but also in promoting equity by enabling educators to design gender-sensitive, personalized instructional interventions based on precise diagnostic profiles rather than global scores. Therefore, this study sought to develop and validate a spatial intelligence test that is not only psychometrically sound but also conducted on 98 English-major university students as a specific target population.

This study sought to answer two research questions:

Research Question One: Does the newly developed spatial intelligence test demonstrate adequate psychometric properties (validity and reliability) for Iranian English-major university students?

Research Question Two: Are there significant gender differences in mastery probabilities of the five spatial intelligence components among Iranian English-major students?

Literature Review

Spatial intelligence refers to the ability to generate, retain, retrieve, and transform well-structured visual images (Lohman, 1996). It enables individuals to perceive relationships among objects, mentally rotate shapes, visualize configurations from different perspectives, and navigate spatial environments (Nagy-Kondor, 2007). Although numerous theoretical frameworks have been proposed to explain spatial ability (e.g., Carroll, 1993; McGee, 1979), the present study draws upon three complementary frameworks. These three were selected because they serve distinct but interrelated purposes: Gardner's theory (1983) broadly positions spatial intelligence as an independent human faculty within multiple intelligences; Thurstone's factor-analytic work (1938) empirically identifies spatial visualization as a primary mental ability; and Mayer's five-factor model (1998, as cited in Nagy-Kondor & Sörös, 2012) provides a fine-grained, operationally measurable structure suitable for diagnostic assessment.

Specifically, Gardner (1983) originally conceptualized spatial intelligence as the ability to perceive and manipulate visual forms, forming one of several independent intelligences (see also Gardner, 2020, for a later memoir). Complementing this, Thurstone (1938) identified spatial visualization as a distinct primary mental ability, defined as the capacity to recognize objects from different angles and reason about spatial relationships (Sorby, 2001). Finally, Mayer (1998, as cited in Nagy-Kondor & Sörös, 2012) decomposed spatial intelligence into five finer components: spatial perception, spatial visualization, mental rotation, spatial relations, and spatial orientation. Although Gardner and Thurstone provide the broader conceptual and empirical rationale for treating spatial ability as a separable dimension, Mayer's five-factor model serves as the primary operational framework for developing the diagnostic test and analyzing gender differences in spatial performance (Nagy-Kondor & Esmailnia, 2021a, 2022). In the contemporary landscape of English for Academic Purposes (EAP), success is increasingly defined not only by linguistic proficiency but also by the underlying cognitive abilities that facilitate academic processing (Hulur et al., 2017).

Menard-Warwick (2008) empirically demonstrated that spatial graphic display enhances English-major readers' comprehension of English sentences with coordinators more effectively than linear sentential representation, indicating that spatial presentation of linguistic information can facilitate learners' processing of complex grammatical structures in academic texts. Furthermore, the spatial organization of texts and the graphical depiction of conceptual relationships have been shown to support learners' ability to recognize structural patterns (Jiang & Grabe, 2007), which is critical for processing dense expository and argumentative materials common in EAP contexts (Carrell, 1985; Grabe, 2009; Jiang & Grabe, 2007). This cognitive processing becomes particularly significant in EAP reading, where learners encounter textbooks in science and engineering that are considered cognitively demanding due to their use of technical terms and specialized grammar (Atai & Hejazi, 2019; Liu et al., 2011; Zand-Moghadam & Khanlarzadeh, 2020). Learners in such contexts often rely on various meaning-making resources to comprehend ideas covered in textbooks, and spatial thinking is considered a malleable skill that can be developed to support STEM learning (National Research Council, 2006; Uttal et al., 2013).

A growing body of research has examined the relationship between spatial intelligence and various dimensions of second language acquisition. In the Iranian English-major context, Molan-Zadeh (2011) investigated the relationship between multiple intelligences and reading comprehension ability among 159 Iranian senior English majors. The findings revealed that spatial intelligence had a statistically significant correlation with participants' reading comprehension scores ($p < .05$), along with intrapersonal, interpersonal, and linguistic intelligences. This suggests that learners with higher spatial ability are better equipped to process visually complex texts, diagrams, tables, and other multimodal representations frequently encountered in academic English reading materials.

In the domain of vocabulary acquisition, Rouhi and Mohebbi (2012) examined the effect of multimedia glosses (pictorial, pictorial + sound, and video) on L2 vocabulary learning among 62 Iranian pre-university students. While their results confirmed the positive effect of multimedia glosses on vocabulary learning, they reported no significant difference between high and low spatial ability groups, suggesting that the relationship between spatial intelligence and vocabulary learning may be mediated by other learner factors or task conditions. Additionally, Zohreh Khorvash & Lotfi (2022) examined spatial encoding in English and Persian, investigating whether typological properties of L1 and L2 affect adult second language acquisition in the spatial domain. The findings demonstrated how cross-linguistic differences in spatial conceptualization influence L2 learners' acquisition of spatial language, including prepositions and locative expressions. This body of research suggests that spatial intelligence plays a differential role across various language skills and components, underscoring the necessity of fine-grained diagnostic assessment.

The intersection of cognitive abilities and gender dynamics presents a complex variable in educational assessment that has received sustained empirical attention (Beikian & Esmailnia, 2023). In English-major contexts, these gender differences have important pedagogical implications. For instance, Burga et al. (2024) investigated gender dynamics in teacher-student interactions during spatial tasks in science classes and found that teachers exhibited prioritized interaction with boys in Year 6, with girls receiving substantially less positive reinforcement and participating less frequently, suggesting that internalized gender stereotypes may influence girls' assertiveness and engagement in spatial activities.

Similarly, Geary et al. (2021) reported that boys' advantages in visuospatial skills ($d_s = .28-.56$) fully mediated sex differences in mathematics achievement, with visuospatial abilities compensating for lower levels of classroom engagement in boys. In the Iranian English-major context, Ghasemi et al. (2025) examined gender differential item functioning of English language items using cognitive diagnostic models, providing methodological insights for detecting gender-based bias in language assessment. In sum, the multidimensional nature of spatial intelligence, as captured by the three complementary frameworks, has been empirically linked to EAP reading comprehension, although its role in vocabulary acquisition remains less clear and likely task-dependent. Critically, gender differences are not uniform across all spatial

sub-skills but appear concentrated in mental rotation – a pattern that is shaped more by environmental and educational opportunities than by innate biological factors. These findings underscore the need for a diagnostic instrument that can provide fine-grained mastery profiles across separate spatial components, which the present study aims to develop and validate for Iranian English-major university students.

Methodology

Design and Participants of the Study

In terms of purpose, the present study is applied, and in terms of nature and method, it is designed within the framework of an exploratory sequential mixed-methods design. This approach enables researchers to first identify the theoretical and structural dimensions of the spatial intelligence construct in the qualitative phase, and subsequently subject the developed instrument to psychometric testing in the quantitative phase. The selection of this design was deemed necessary because the precise measurement of spatial intelligence dimensions necessitates the simultaneous utilization of experts' theoretical perspectives and quantitative psychometric analyses. Accordingly, this research has proceeded step-by-step from the extraction of theoretical dimensions to the testing of empirical hypotheses.

The present study employed a three-phase participant selection strategy aligned with its exploratory sequential mixed-methods design. Participants were recruited across qualitative, quantitative, and validation phases to ensure comprehensive instrument development and psychometric evaluation.

Phase 1: Qualitative Expert Panel: In the initial qualitative phase, a purposive sample of ten international experts in spatial intelligence and cognitive psychometrics was recruited and interviewed to inform the theoretical framework and item development of the spatial intelligence test. These experts were affiliated with prestigious institutions worldwide, including Northwestern University, University of Debrecen, University of California, and other leading research centers. The semi-structured interviews (conducted via Zoom and email) drew upon extensive publication records and expertise in spatial cognition, test construction, and cognitive diagnostic modeling ensured that the theoretical underpinnings and content validity of the instrument met contemporary psychometric standards.

Ten international experts in spatial intelligence and cognitive psychometrics were recruited and interviewed (via Zoom/email) to inform the test's framework and item development. Their expertise and publications ensured the content validity of the instrument.

Phase 2: Quantitative Sample for Psychometric Evaluation: In the quantitative phase, the participants were 250 high school students selected from a target population of approximately 8,500 students enrolled in 75 public schools (38 girls' schools and 37 boys' schools) in Qaemshahr County, Mazandaran Province, Iran. Table 1 presents the demographic characteristics of the 250 participants in the quantitative phase.

Table 1

Demographic Characteristics of the Quantitative Sample (N = 250)

Demographic Variable	Category	N	Percent (%)
Gender	Male	128	51.2
	Female	122	48.8
	Total	250	100.0
Grade Level	10th Grade	85	34.0
	11th Grade	88	35.2
	12th Grade	77	30.8

	Total	250	100.0
School Type	Public	250	100.0
	Private	–	–
	Total	250	100.0

Phase 3: Validation Sample in English-major Context

To examine the psychometric properties of the test for English major contexts, a final validation phase was conducted with 98 undergraduate students majoring in English Language Teaching or English Literature at universities in Mazandaran Province, Iran. Participants were selected using convenience and purposive sampling based on three inclusion criteria: (1) enrollment in an English major, (2) voluntary informed consent, and (3) physical presence during testing. Although females typically outnumber males in these majors in Iran, the present sample comprised 52 males and 46 females – a slight male majority due to the specific accessibility of participants in the sampled classes. The sample size was determined based on two considerations. First, a power analysis using GPower software (Faul et al., 2009) indicated that a minimum of 85 participants was required to achieve 80% statistical power for detecting medium effect sizes in correlation and regression analyses ($\alpha = .05$). Second, for cognitive diagnostic modeling (CDM), the sample size of 98 is within the range of sample sizes ($N = 100\text{--}200$) commonly used in CDM simulation studies that have demonstrated acceptable parameter recovery and classification accuracy (e.g., Kunina-Habenicht et al., 2012; Sen & Cohen, 2021). Thus, the sample size of 98 is deemed adequate for the CDM analyses reported in this study.

The initial standardization of the test was carried out on a general high school population (Phase 2, $N = 250$). For the final validation, the test was re-administered to the specific target population: English-major university students. A convenience and purposive sampling procedure was employed for this phase. The frequency distribution of participants by academic year and gender is presented in Table 2.

Table 2

Demographic Characteristics of the Validation Sample ($N = 98$)

Students	Gender	N	Percent
Freshman (first-year)	Male	17	17.3
Freshman (first-year)	Female	14	14.3
Sophomore	Male	13	13.3
Sophomore	Female	12	12.2
Junior	Male	12	12.2
Junior	Female	10	10.2
Senior	Male	11	11.2
Senior	Female	9	9.2
Total		98	100.0

As shown in Table 2, the sample demonstrated adequate diversity in terms of gender distribution and academic year representation, which enhances the representativeness of the sample relative to the target population of English language students in Mazandaran Province. This balanced distribution across academic levels (first- to fourth-year students) and gender groups supports the generalizability of the psychometric findings and allows for more robust subgroup analyses in future research examining the relationship between spatial intelligence and language learning outcomes.

Instruments

The primary instrument employed in this study was a teacher-made test, which was designed and psychometrically validated for the first time within the Iranian educational context. This instrument comprises 25 multiple-choice items constructed based on findings from the qualitative phase and aligned with the five core dimensions of spatial intelligence. While Gardner's (1983) and Thurstone's (1938) frameworks provided the broad conceptual rationale for treating spatial ability as a multidimensional construct, Mayer's (1998) five-factor model served as the direct operational framework for item development, defining the specific sub-skills measured by the test:

1. Mental Rotation (7 items; e.g., rotating a shape 180°, counting hidden cubes, selecting rotated projections): The ability to mentally manipulate and rotate two- or three-dimensional objects to match a target configuration.
2. Spatial Visualization (7 items; e.g., combining shapes, completing dot-position sequences, mirror images): The capacity to analyze, decompose, and reconstruct complex spatial relationships among components of a figure or scene.
3. Understanding of 3D Relationships (4 items; e.g., folding nets into cubes, mirror reflections, 3D pattern continuation): The skill to perceive and reason about the position, orientation, and structure of objects within a three-dimensional space.
4. Spatial Perceptual Speed (4 items, timed; e.g., rapid pattern continuation, embedded figures): The ability to rapidly and accurately identify patterns, similarities, differences, and spatial positions among visual stimuli.
5. Visual Memory (3 items; e.g., counting cubes after brief viewing, continuing visual sequences without delay): The capacity to encode, store, and retrieve visual-spatial information and mental images.

To ensure the quality of the instrument, multiple levels of validity were assessed, including content validity (using CVR and CVI indices), construct validity (through exploratory and confirmatory factor analysis), and convergent and discriminant validity (using the AVE index). Construct validity was measurable because each of the five spatial components was operationally defined prior to factor analysis (see operational definitions above). The reliability of the instrument was also calculated using Cronbach's alpha and composite reliability methods.

Data Collection Procedure

Qualitative Phase

In the qualitative phase, data collection was conducted through semi-structured online interviews with international experts in spatial intelligence and psychometrics via Zoom, complemented by content evaluation of test items. The expert panel included distinguished scholars from prestigious institutions such as Northwestern University, University of Debrecen, and University of California, alongside recognized domestic experts in the field. To ensure comprehensiveness of the data, a systematic content analysis of credible scientific sources from both internal and external databases was also conducted. To ensure the trustworthiness of qualitative findings, the criteria proposed by Guba and Lincoln (1994). Trustworthiness was ensured by: member checking (credibility), thick description (transferability), audit trail (dependability), and reflexivity (confirmability) (Danaiifard & Mozaffari, 2008, as cited in Yekta & Shafieabadi, 2021).

Quantitative Phase

In the quantitative phase, the teacher-made test (25 items) was administered to a sample of 250 high school students (128 boys, 122 girls) in Qaemshahr, with consideration of gender and geographic diversity. The test items were designed based on the five dimensions of Mayer's (1998) model, which had been confirmed and operationalized through the qualitative phase (expert interviews and content analysis of the literature). A multi-stage cluster sampling procedure was employed to ensure adequate representation of girls' and boys' schools [Stage 1: Geographic stratification; Stage 2: Selection of primary clusters (districts); Stage 3: Identification and selection of secondary clusters (schools); Stage 4: Field visits and selection of the final sample (students)]. With the operational definitions of the five components provided in the Instruments section, construct validity could be meaningfully assessed. Thus, construct validity was examined through EFA and CFA by examining fit indices such as RMSEA, CFI, TLI, GFI, and SRMR. Convergent and divergent validity were estimated using AVE and the Fornell-Larcker criterion, respectively. For concurrent validity, we initially used the DAT-SR, acknowledging that it primarily measures spatial relations and visualization, not the full five-component model. Therefore, the obtained correlation ($r = 0.76$) should be interpreted as partial concurrent validity; a more comprehensive battery would be desirable in future studies. The primary validity evidence for the full five-factor structure comes from EFA/CFA (see Results).

Data Analysis

Quantitative data were analyzed using SPSS, AMOS, and SmartPLS software. In addition to descriptive statistics, advanced inferential analyses such as structural equation modeling and cognitive diagnostic models (CDM) were employed to obtain students' cognitive mastery profiles for each of the five components. Specifically, the G-DINA model estimated the probability that each student mastered each of the five spatial skills (e.g., 0.77 for mental rotation). Then, these probabilities were converted into simple mastery classifications (mastery = above 0.70, non-mastery = 0.70 or below). To compare male and female performance at the skill level, chi-square tests were conducted on these classifications (e.g., 77% of males vs. 63% of females mastered mental rotation, $p < .05$). Thus, the integration of qualitative and quantitative findings not only facilitated the development of a valid and localized instrument, but also provided the foundation for offering evidence-based educational recommendations and designing gender-sensitive programs in Iranian schools and universities. Cognitive Diagnostic Model (CDM) Specification: The G-DINA model was fitted to the data using the CDM package (version 7.5; George et al., 2016) in R. The Q-matrix was developed based on expert panel judgment.

Results

In this section, the findings of the present study are presented and examined. The first part describes the standardization procedure of the questionnaire, followed by an explanation of the evaluation process administered to English language students:

To calculate the Content Validity Ratio (CVR), ten experts in the field of cognitive psychometrics and spatial intelligence were asked to evaluate each of the 25 test items based on a three-point Likert scale (1 = not necessary, 2 = useful but not necessary, 3 = necessary). According to Lawshe's (1975) table, for a panel of 10 experts, the minimum acceptable CVR value is 0.62. The results indicated that all CVR values ranged from 0.80 to 1.00, exceeding the required threshold.

$$CVR = \frac{n_e - \frac{N}{\gamma}}{\frac{N}{\gamma}}$$

To calculate the Content Validity Index (CVI) for face validity, the panel of experts evaluated each item across three dimensions—clarity, relevance, and usefulness—using a four-point Likert scale (1 = inappropriate, 2 = somewhat appropriate, 3 = appropriate, 4 = completely appropriate). Based on Lynn's (1986) criterion, the minimum acceptable threshold for CVI is 0.79. The findings indicated that all CVI values across the three aforementioned dimensions ranged from 0.80 to 1.00, exceeding the required threshold.

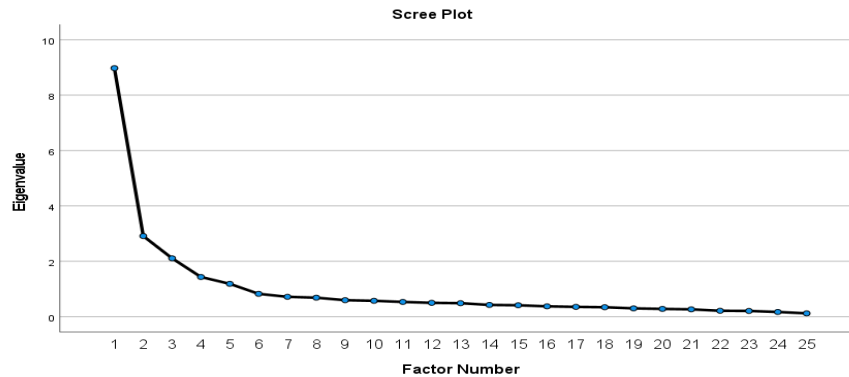
The CVI was calculated using the following formula:

$$CVI = \left(\frac{\text{Number of experts rating the item as 3 or 4}}{\text{Total number of experts}} \right)$$

In the next step, construct validity was examined. The results of exploratory factor analysis (EFA) conducted using SPSS software indicated that the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.91, which exceeds the minimum threshold of 0.50. Therefore, the sample size was deemed sufficient for factor analysis. Additionally, Bartlett's test of sphericity was statistically significant ($p = .000$, $\chi^2 = 4382.2$), indicating that the correlations among variables were significantly different from zero.

Figure 1

Scree plot of eigenvalues for the 25 items



The figure shows a clear elbow after the fifth factor, supporting the extraction of five components. The results of exploratory factor analysis conducted on the 25 items with Varimax rotation indicated that, based on the Kaiser criterion (eigenvalues greater than 1) presented in Table 3 and the scree plot shown in Figure 1, the identification of five components was logical, collectively explaining 59.18% of the total variance.

Table 3

Eigenvalues

Factor	Total	% of Variance	Cumulative %
1	4.20	16.80	16.80
2	4.13	16.54	33.34
3	3.02	12.11	45.46
4	1.84	7.38	52.84
5	1.58	6.34	59.18

Table 4 presents the rotated factor loading matrix. Items with high loadings on a single factor formed five distinct factors. Factor 1 (Mental Rotation) included items 1, 3, 5, 6, 7, 8, and 22; Factor 2 (Spatial Visualization) included items 4, 9, 10, 11, 12, 14, and 25; Factor 3 (Visual Memory) included items 15, 16, 17, 18, 19, and 20; Factor 4 (Spatial Perceptual Speed) included items 13, 21, and 23; and Factor 5 (Understanding 3D Relationships) included items 2 and 24. All items demonstrated factor loadings above 0.40. Therefore, these items were identified as components of the spatial intelligence test.

Table 4

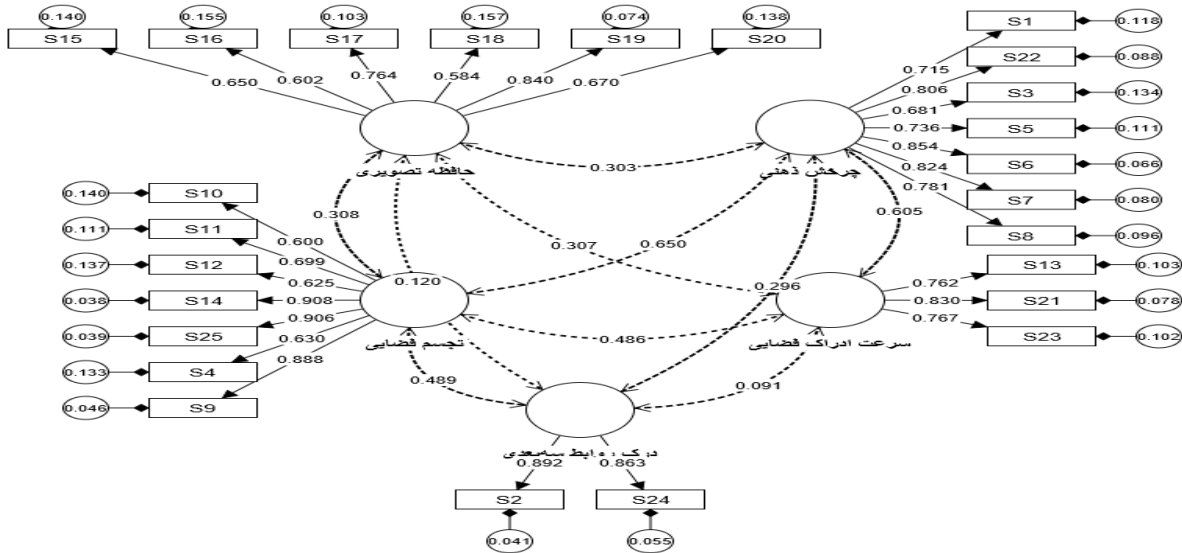
Factor Loadings for Each Item

Item	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
22	.775	–	–	–	–
7	.763	–	–	–	–
6	.759	–	–	–	–
8	.706	–	–	–	–
5	.656	–	–	–	–
1	.618	–	–	–	–
3	.581	–	–	–	–
14	–	.843	–	–	–
25	–	.796	–	–	–
9	–	.771	–	–	–
11	–	.668	–	–	–
4	–	.600	–	–	–
12	–	.551	–	–	–
10	–	.530	–	–	–
19	–	–	.883	–	–
17	–	–	.747	–	–
20	–	–	.659	–	–
15	–	–	.642	–	–
16	–	–	.600	–	–
18	–	–	.553	–	–
23	–	–	–	.739	–
13	–	–	–	.688	–
21	–	–	–	.672	–
2	–	–	–	–	.889
24	–	–	–	–	.770

The results of confirmatory factor analysis (CFA) conducted using SmartPLS software, as illustrated in Figures 2 and 3, indicated that the factor loadings for all items exceeded 0.50, demonstrating that all items possess adequate explanatory power. Furthermore, the absolute T-values for all items were greater than 1.96, indicating that each item significantly measures its corresponding latent variable.

Figure 2

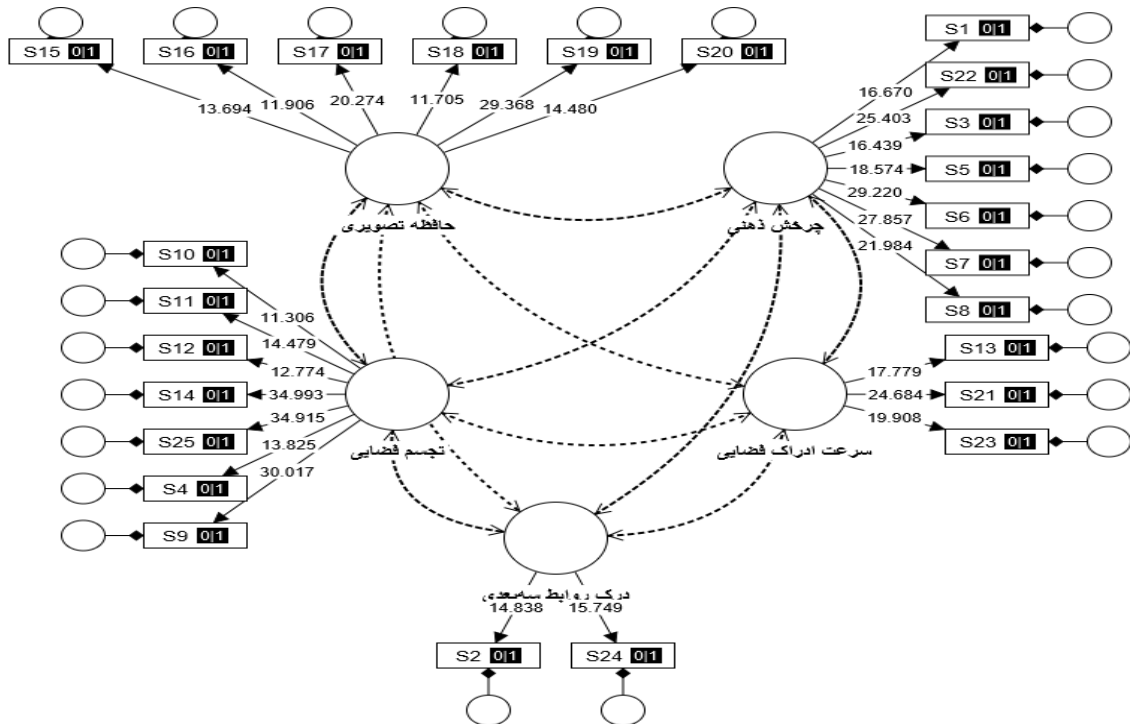
Factor Loadings of The 25 Items on Each of the Five Spatial Components



All loadings exceed 0.50, indicating adequate item-factor relationships.

Figure 3

T-values for the factor loadings



All absolute T-values exceed 1.96, confirming that each item significantly measures its corresponding latent component at $p < .05$. Table 5 presents the goodness-of-fit indices, including RMSEA, Chi-square, NNFI, CFI, GFI, and SRMR. As can be observed, all indices fall within acceptable ranges, and the confirmatory factor analysis supports the five-factor structure. Therefore, the spatial intelligence test demonstrates acceptable construct validity.

Table 5*Goodness-of-Fit Indices*

Index	Value	Acceptable Criterion	Interpretation
Chi-square (χ^2)	$\chi^2 = 442.1, p = .000$	Should be significant	$p < .05$, therefore significant
χ^2/df	1.66	Values between 1 and 3	Acceptable
RMSEA	0.047	Values less than 0.08	Acceptable
NFI	0.90	Values between 0.90 and 0.95	Acceptable
TLI	0.95	Values between 0.90 and 0.95	Acceptable
CFI	0.95	Values between 0.90 and 0.95	Acceptable
GFI	0.89	Values between 0.90 and 0.95	Acceptable
SRMR	0.046	Below 0.08	Acceptable

In this section, the results of administering the spatial intelligence test to English language students at the university are presented: The reliability of the test was calculated using internal consistency (Cronbach's alpha and composite reliability) among English language students. The results of the measurement model analysis using SmartPLS software, presented in Table 6, indicated that Cronbach's alpha values were 0.88 for spatial visualization, 0.85 for visual memory, 0.90 for understanding of 3D relationships, 0.85 for spatial perceptual speed, and 0.90 for mental rotation. Furthermore, composite reliability values were 0.90 for spatial visualization, 0.85 for visual memory, 0.85 for understanding of 3D relationships, 0.91 for spatial perceptual speed, and 0.90 for mental rotation. All these indices exceeded the acceptable threshold of 0.70. Therefore, the internal consistency (reliability) of items for each component is satisfactory.

Table 6*Cronbach's Alpha and Composite Reliability*

variable	Cronbach's Alpha	Composite Reliability
Spatial Visualization	0.880	0.903
Visual Memory	0.851	0.856
Understanding of 3D Relationships	0.902	0.850
Spatial Perceptual Speed	0.850	0.914
Mental Rotation	0.902	0.907

In the final stage, gender was examined based on the percentage of mastery over the dimensions of spatial intelligence among English university students. Data analysis revealed that gender differences in spatial intelligence were significant only in three components. Male students demonstrated significantly better performance ($p < .05$) in Mental Rotation (.77 vs. .63), Understanding of 3D Relationships (.64 vs. .51), and Spatial Visualization (.74 vs. .63). In the remaining components—including Spatial Perceptual Speed (.69 vs. .66) and Visual Memory (.75 vs. .72)—no significant differences were observed ($p > .05$). Furthermore, the overall spatial intelligence score was significantly higher for male students (71.8) compared to female students (63.0) ($p < .05$). These findings indicate that gender differences in spatial intelligence are not universal but are concentrated in specific dimensions.

Table 7*Probability of Skill Mastery Among English-major University Students (N = 98)*

Skill	Female Mastery (%)	Male Mastery (%)	Difference	χ^2 (df = 1)	Chi-Square Result	Test
Mental Rotation	63	77	-14	4.21	p < .05	
Spatial Visualization	63	74	-11	2.10	p > .05	
Understanding of 3D Relationships	51	64	-13	3.89	p < .05	
Spatial Perceptual Speed	66	69	-3	0.21	p > .05	
Visual Memory	72	75	-3	0.29	p > .05	
Overall Spatial Intelligence Score	63.0	71.8	-8.8	4.52	p < .05	

In mental rotation, the mastery percentage of male students (77%) was significantly higher than that of female students (63%) ($p < .05$). This 14-percentage-point difference in mental rotation ability is consistent with the findings of numerous previous studies reporting that males and boys tend to perform better on this dimension. In spatial visualization, a difference of 11 percentage points was observed between male (74%) and female (63%) students; however, this difference did not reach statistical significance ($p > .05$). This result indicates that male and female students demonstrated different performance patterns on this component. Regarding understanding of 3D relationships, male students showed a higher mastery percentage (64%) compared to female students (51%). The 13-percentage-point difference was statistically significant ($p < .05$), suggesting that this component may be particularly sensitive to potential gender differences or variations in spatial experience.

In spatial perceptual speed, the difference between the two groups was negligible and non-significant ($p > .05$). This component, which pertains to the rapid identification of patterns and spatial positions, appears to be less influenced by gender. Similarly, in visual memory, the mastery percentages of both groups were nearly equivalent, and the difference was non-significant ($p > .05$). These findings suggest that visual memory capacity is comparable across both gender groups. Overall, the findings presented in Table 7 indicate that gender differences were primarily significant in three components—mental rotation, spatial visualization, and understanding of 3D relationships—with male students demonstrating superior performance in these three skills. In the remaining components (spatial perceptual speed and visual memory), no significant differences were observed.

Discussion

This study aimed to develop and validate a spatial intelligence test based on cognitive diagnostic models and to examine gender differences in mastery probabilities of five spatial components among Iranian English-major university students. The findings provided strong psychometric support for the newly developed instrument and revealed that gender differences were not uniform across all spatial dimensions but were concentrated in mental rotation, spatial visualization, and understanding of 3D relationships. These results are discussed in light of previous theoretical and empirical literature, followed by pedagogical implications, limitations, and future research directions.

Psychometric Properties of the Spatial Intelligence Test

Regarding the first research question, the findings indicate that the newly developed spatial intelligence test possesses satisfactory psychometric properties for Iranian English-major university students. As reported in the Results section, content validity indices (CVR and CVI) exceeded the recommended thresholds (Lawshe, 1975; Lynn, 1986), and confirmatory factor analysis confirmed a five-factor structure consistent with Mayer's (1998) model, with fit indices ($\chi^2/df = 1.66$, RMSEA = 0.047, CFI = 0.95, TLI = 0.95, SRMR

= 0.046) within conventionally acceptable ranges (Hu & Bentler, 1999; Kline, 2015). Cronbach's alpha values (0.85–0.90) suggested acceptable internal consistency. These psychometric properties are a necessary prerequisite for a diagnostic instrument intended for educational settings. Moreover, the successful application of the G-DINA model demonstrates that cognitive diagnostic modeling is feasible for analyzing spatial ability data in this population, allowing for the fine-grained mastery profiles that were previously unavailable. Thus, the test meets the basic validity and reliability standards required for further research and potential classroom application.

Gender Differences in Spatial Intelligence Components

Regarding the second research question, the results (see Table 7) revealed that gender differences were not uniform across all spatial components. Male participants outperformed female participants in three higher-order spatial manipulation skills: mental rotation, spatial visualization, and understanding of 3D relationships. In contrast, no significant gender differences were observed for spatial perceptual speed or visual memory. This pattern suggests that gender gaps in spatial intelligence are concentrated in specific dimensions rather than being a general phenomenon.

The male advantage in mental rotation is highly consistent with a large body of international research. For instance, Wu et al. (2025) reported that primary school boys outperformed girls in mental rotation tasks using behavioral and fNIRS measures, and a recent meta-analysis by Xue et al. (2025) confirmed that male advantages in spatial navigation (which involves mental rotation) are robust across ages and task conditions. Similarly, Geary et al. (2021) found that boys' superior visuospatial skills mediated gender differences in mathematics achievement. In the Iranian English-major context, the present finding corroborates the descriptive patterns observed by Nagy-Kondor and Esmailnia (2021a, 2022), although their studies did not focus on language learners specifically. The 14-percentage-point difference in mental rotation mastery in the current study is substantial and suggests that this dimension may be particularly sensitive to gender-related factors such as differential practice with spatial games, educational experiences, and possibly stereotype threat (Ebert et al., 2025; Fioriti et al., 2024).

The significant difference in spatial visualization (11% advantage for males) also aligns with previous studies. Spatial visualization involves more complex mental manipulation than simple rotation, including decomposition and reconstruction of spatial configurations. This finding is consistent with the work of Uttal et al. (2013), who showed that spatial visualization is malleable but often exhibits male advantages in untrained populations. In English-major reading comprehension, Molan-Zadeh (2011) reported a positive correlation between spatial intelligence and reading scores. Consistent with this, the present study found a gender difference in spatial visualization (74% vs. 63%). This correlational finding suggests a possible association between gender and performance in spatial visualization, but it does not imply that being female directly causes cognitive challenges. Future experimental research would be needed to examine causality.

Understanding of 3D relationships showed the largest relative male advantage (13 percentage points). This component requires reasoning about the orientation and structure of three-dimensional objects. This result is consistent with the findings of Nagy-Kondor and Esmailnia (2021b), who demonstrated that engineering students' performance on 3D tasks differed by gender. In the EAP context, where science and engineering textbooks frequently contain three-dimensional illustrations, this observed gender gap may be associated with differences in performance on 3D tasks. However, because the study only measured mastery probabilities (correlational data), no causal claim can be made about gender leading to unequal access. Instead, we note that if replicated in experimental studies, such a gap might warrant compensatory instructional scaffolding (Atai & Hejazi, 2019; Liu et al., 2011).

In contrast, the absence of significant gender differences in spatial perceptual speed and visual memory is noteworthy. Spatial perceptual speed involves rapid identification of patterns and differences—

a more automatic, less strategically controlled process. Visual memory, the ability to encode and retrieve visual-spatial information, also showed no gender gap. These findings suggest that not all spatial abilities are equally susceptible to gender influences. They are consistent with some meta-analyses that report small or null gender effects on perceptual speed (Linn & Petersen, 1985) and visual memory (Cowan, 2017). In English-major classrooms, this implies that male and female learners are similarly capable of rapidly recognizing visual patterns and remembering spatial information, which can be leveraged for equitable instruction.

Taken together, the pattern of results—significant differences in three higher-order spatial manipulation components but not in lower-level perceptual or memory components—supports the view that gender differences in spatial intelligence are task-specific and mediated by environmental and educational factors rather than being purely biological (Ebert et al., 2025; Fioriti et al., 2024). In the Iranian English-major context, where school experiences and extracurricular activities may differ by gender, these findings underscore the need for diagnostic-based instructional interventions.

Implications for English-major Pedagogy and Assessment

The findings carry several implications for English language teaching and assessment in academic contexts. First, the validated spatial intelligence test can serve as a diagnostic tool for EAP instructors to identify learners' cognitive profiles. For example, female students who show lower mastery in mental rotation and 3D relationships could benefit from explicit training in interpreting spatial diagrams, flowcharts, and graphical abstracts commonly found in STEM-related English texts (National Research Council, 2006; Uttal et al., 2013). Second, materials developers may wish to integrate spatial skill-building activities into EAP curricula, such as mental rotation exercises using 3D models or visualization tasks that accompany reading passages. Third, assessment practices in EAP could move beyond unitary language scores to include cognitive diagnostic feedback, enabling personalized learning pathways (De la Torre & Sorrel, 2023).

One possible explanation for the persistent male advantage in mental rotation and 3D relationships in the present sample is the differential exposure to spatial activities during childhood and adolescence. In the Iranian educational context, extracurricular opportunities (e.g., construction toys, video games, and hands-on technical activities) are more commonly available to boys than to girls, which may lead to differential practice effects. Additionally, stereotype threat – the anxiety that one might confirm a negative stereotype about one's group – could have contributed to lower performance among female participants, as suggested by previous research (Ebert et al., 2025; Fioriti et al., 2024).

Conclusions

In summary, the psychometric analyses (see Results) confirmed that the newly developed 25-item test is a valid and reliable instrument for assessing spatial intelligence among Iranian English-major university students. Regarding gender differences, the findings indicated that male participants outperformed female participants in three higher-order spatial components (mental rotation, spatial visualization, and understanding of 3D relationships), but not in perceptual speed or visual memory. These results are discussed in detail above. The following practical implications, limitations, and suggestions for future research are therefore offered.

The findings have several practical implications for EAP pedagogy and assessment. English-major instructors may use diagnostic profiles to design targeted interventions, such as flowchart-based mental rotation tasks or annotated 3D diagrams for learners struggling with STEM texts. Syllabus designers may wish to integrate spatial skill-building activities, including graphic organizers and visual mapping exercises, into language courses. Assessment developers may adopt cognitive diagnostic models to

provide visual feedback reports (e.g., radar charts or color-coded bar graphs) that pinpoint specific spatial weaknesses beyond unitary test scores.

Several limitations warrant consideration. The validation sample appears limited to 98 English-major students from one Iranian province, which may restrict generalizability. No standardized English proficiency measure seems not to have been administered, so the interaction between spatial ability and L2 achievement might not be fully examined. Moreover, the Q-matrix appears to have been developed based on expert judgment alone; future studies should validate it using empirical response data or think-aloud protocols. The cross-sectional design also may preclude causal conclusions regarding the origins of gender differences.

Future research might extend the validation of the test to larger and more diverse samples of English-major students across different regions and educational levels. Longitudinal studies may be needed to examine whether explicit spatial training might reduce gender gaps and improve reading comprehension of visually complex academic texts. Additionally, integrating the spatial intelligence test with other cognitive measures (e.g., working memory, language aptitude) might provide a more comprehensive understanding of individual differences in EAP success.

In conclusion, this study may contribute a psychometrically sound, diagnostically informative, and culturally adapted spatial intelligence test for Iranian English-major university students. By revealing dimension-specific gender differences, the findings appear to underscore the need for personalized, evidence-based instructional approaches that support equitable learning outcomes in English for Academic Purposes. Specifically, the diagnostic profiles generated by this test could enable educators to identify whether a learner's difficulty with STEM-related reading materials stems from weak mental rotation, poor spatial visualization, or limited understanding of 3D relationships. Such fine-grained information may then guide targeted interventions – such as visualization training or diagram-based reading strategies – without over-generalizing based on gender. Hence, the test may serve not only as a research instrument but also as a practical tool for promoting fairness and individualized support in EAP classrooms.

Acknowledgment

The authors are deeply grateful to all participants who generously volunteered their time and cooperated throughout the data gathering phase.

Declaration of Conflicting Interests

The authors hereby declare that they have no competing interests or conflicts of interest relevant to this study.

Funding Details

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Atai, M. R., & Hejazi, S. Y. (2019). Assessment of academic English language needs of Iranian post-graduate students of psychology. *Ibérica*, (38), 275-302. <https://doi.org/10.17398/2340-2784.38.275>

- Beikian, A., & Esmailnia, S. (2023). Effectiveness of data-driven learning in reducing phrasal verb-related errors among intermediate English as a foreign language learners. *Iranian Journal of English for Academic Purposes*, 12(1), 1-18. https://ijeap.cmu.ac.ir/article_169561.html
- Burga, E. G., Tenbrink, T., & Mills, D. (2024, March). Gender in teacher-student interactions: Another factor in spatial ability development and STEM affiliation. In *Spatial Cognition Conference Proceedings 2024*. <https://doi.org/10.1007/978-3-031-63115-3>
- Carrell, P. L. (1985). Facilitating ESL reading by teaching text structure. *TESOL quarterly*, 19(4), 727-752. <https://doi.org/10.2307/3586673>
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies* (Vol. 1). Cambridge university press. <https://doi.org/10.1017/CBO9780511571312>
- Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic bulletin & review*, 24(4), 1158-1170. <https://doi.org/10.3758/s13423-016-1191-6>
- Dahmardeh, M., Mohammed Hasan, A., Mohammadi, P., & Al-Rashdi, F. (2025). A comparative study of gender representation in EFL coursebooks in the Middle East. *Journal for Multicultural Education*, 19(1), 108-120. <https://doi.org/10.1108/jme-10-2024-0138>
- De La Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of educational and behavioral statistics*, 34(1), 115-130. <https://doi.org/10.3102/1076998607309474>
- De la Torre, J., & Sorrel, M. A. (2023). Cognitive diagnosis models. *New handbook of mathematical psychology*, 3, 385-420. <https://doi.org/10.1017/9781108902724.010>
- Ebert, W. M., Jost, L., & Jansen, P. (2025). No relationship between gender stereotypes and mental rotation in preschool girls. *Frontiers in Psychology*, 16, 1650979. <https://doi.org/10.3389/fpsyg.2025.1650979>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149– 1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fioriti, C. M., Martell, R. N., Daker, R. J., Malone, E. P., Sokolowski, H. M., Green, A. E., ... & Lyons, I. M. (2024). Examining the interplay between the cognitive and emotional aspects of gender differences in spatial processing. *Journal of Intelligence*, 12(3), 30. <https://doi.org/10.3390/jintelligence12030030>
- Gardner, H. (2020). *A synthesizing mind: A memoir from the creator of multiple intelligences theory*. MIT Press. <https://doi.org/10.7551/mitpress/12405.001.0001>
- Geary, D. C., Hoard, M. K., & Nugent, L. (2021). Boys' visuospatial abilities compensate for their relatively poor in-class attentive behavior in learning mathematics. *Journal of experimental child psychology*, 211, 105222. <https://doi.org/10.1016/j.jecp.2021.105222>
- Ghasemi, S., Delavar, A., & Eskandari, F. (2025). Evaluating of gender differences in English language items on the master of psychology series exam using the wald test, likelihood ratio test, and revised likelihood ratio test based on the cognitive diagnostic model. *Educational Measurement*, 16(59), 37-61. <https://doi.org/10.22054/jem.2024.73600.3460>
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge university press. <https://doi.org/10.1017/CBO9781139150484>

- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, 2(163-194), 105. https://miguelangelmartinez.net/IMG/pdf/1994_Guba_Lincoln_Paradigms_Quali_Research_chapter.pdf
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Hülür, G., Ram, N., Willis, S. L., Schaie, K. W., & Gerstorff, D. (2016). Cognitive aging in the Seattle longitudinal study: Within-person associations of primary mental abilities with psychomotor speed and cognitive flexibility. *Journal of Intelligence*, 4(3), 12. 10.3390/jintelligence4030012
- Jiang, X., & Grabe, W. (2007). Graphic organizers in reading instruction: Research findings and issues. *Reading in a foreign language*, 19(1), 34-55. <https://doi.org/10.64152/10125/66616>
- Khorvash, Z., & Lotfi, A. (2022). Spatial encoding in English and Persian: Typological influences on second language acquisition. *International Journal of Foreign Language Teaching and Research*, 10(41), 49-63. <https://doi.org/10.30495/jfl.2022.689309>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Kunina-Habenicht, O., Rupp, A. A., & Wilhelm, O. (2012). The impact of model misspecification on parameter estimation and item-fit assessment in log-linear diagnostic classification models. *Journal of Educational Measurement*, 49(1), 59-81. <https://doi.org/10.1111/j.1745-3984.2011.00160.x>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child development*, 1479-1498. <https://doi.org/10.2307/1130467>
- Liu, J. Y., Chang, Y. J., Yang, F. Y., & Sun, Y. C. (2011). Is what I need what I want? Reconceptualising college students' needs in English courses for general and specific/academic purposes. *Journal of English for Academic Purposes*, 10(4), 271-280. <https://doi.org/10.1016/j.jeap.2011.09.002>
- Lohman, D. F. (1996). Spatial ability and g. *Human abilities: Their nature and measurement*, 97(116), 1.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382-385. <https://doi.org/10.1097/00006199-198611000-00017>
- Mahmoudikia, M., & Ahmadi, A. (2020). Iranian university instructors' perception on English-majoring students' speaking needs: A case of Shiraz university English department. *Iranian Journal of English for Academic Purposes*, 9(4), 15-30. https://ijeap.cmu.ac.ir/article_117527.html
- McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological bulletin*, 86(5), 889. <https://doi.org/10.1037/0033-2909.86.5.889>
- Menard-Warwick, J. U. L. I. A. (2008). The cultural and intercultural identities of transnational English teachers: Two case studies from the Americas. *TESOL quarterly*, 42(4), 617-640. <https://doi.org/10.1002/j.1545-7249.2008.tb00151.x>
- Molan-Zadeh, F. (2011). *The relationship between Iranian EFL learners' multiple intelligences and their reading comprehension ability* [Master 's thesis, Payame Noor University].

- Mpiti, V. S., Ncanywa, T., & Asaleye, A. J. (2025). Do educators' demographic characteristics drive learner academic performance? Examining the role of gender, qualifications, and experience. *Education Sciences, 15*(4), 487. <https://doi.org/10.3390/educsci15040487>
- Nagy-Kondor, R. (2007, January). Spatial ability of engineering students. In *Annales Mathematicae et Informaticae* (Vol. 34, No. 2007, pp. 113-122). https://www.researchgate.net/publication/228611356_Spatial_ability_of_engineering_students
- Nagy-Kondor, R., & Esmailnia, S. (2021 a). A comparison between the spatial intelligence of the freshman engineering students of Sharif University of Technology and Debrecen University. *Periodica Polytechnica Social and Management Sciences, 29* (2), 159-167. <https://doi.org/10.3311/ppso.15740>
- Nagy-Kondor, R., & Esmailnia, S. (2021b). Polyhedrons vs. curved surfaces with mental cutting: Impact of spatial ability. *Acta Polytechnica Hungarica, 18*(6), 71–83. <https://doi.org/10.12700/APH.18.6.2021.6.4>
- Nagy-Kondor, R., & Esmailnia, S. (2022). Development of spatial ability extra tasks (SAET): Problem solving with spatial intelligence. *Quality & Quantity, 56* (5), 3751-3768. <https://doi.org/10.1007/s11135-021-01284-7>
- Nagy-Kondor, R., & Sorous, M. (2012). *Spatial abilities in engineering education*. University of Debrecen Press.
- National Research Council, Division on Earth, Life Studies, Board on Earth Sciences, Geographical Sciences Committee, Committee on Support for Thinking Spatially, & The Incorporation of Geographic Information Science Across the K-12 Curriculum. (2005). *Learning to think spatially*. National Academies Press. <https://doi.org/10.17226/21881>
- Ouyang, X., Zhang, X., & Zhang, Q. (2022). Spatial skills and number skills in preschool children: The moderating role of spatial anxiety. *Cognition, 225*, 105165. <https://doi.org/10.1016/j.cognition.2022.105165>
- Rouhi, A., & Mohebbi, H. (2013). Glosses, spatial intelligence, and L2 vocabulary learning in multimedia context. *3L: The Southeast Asian Journal of Language Studies, 19*(2), 75-87.
- Rupp, A. A., & Templin, J. L. (2007, April). Unique characteristics of cognitive diagnosis models. In *annual meeting of the National Council on Measurement in Education, Chicago*.
- Sen, S., & Cohen, A. S. (2021). Sample size requirements for applying diagnostic classification models. *Frontiers in Psychology, 11*, 621251. <https://doi.org/10.3389/fpsyg.2020.621251>
- Sorby, S. A. (1999). Developing 3-D spatial visualization skills. *The Engineering Design Graphics Journal, 63*(2).
- Suprpto, P. K., Bin Ahmad, M. Z., Chaidir, D. M., Ardiansyah, R., & Diella, D. (2018). Spatial intelligence and students achievement to support creativity on visuospatial-based learning. *Jurnal Pendidikan IPA Indonesia, 7*(2), 224-231. <https://doi.org/10.15294/jpii.v7i2.14322>
- Tahir, Z., Saleh, R. R., Fadhil, S., Awat, C., & Ibrahim, S. (2025). Pedagogical training in Kurdistan: A critical analysis of present challenges. *Amandemen: Journal of Learning, Teaching and Educational Studies, 3*(1), 117-134. <https://doi.org/10.61166/amd.v3i1.97>
- Thurstone, L. L. (1938). *Primary mental abilities* (Vol. 1). University of Chicago Press. https://pure.mpg.de/rest/items/item_2398946_5/component/file_2603779/content

- 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. <https://doi.org/10.1037/a0028446>
- Wafa, M. N., Zia, Z., & Frozan, F. (2023). Consistency and ability of students using DINA and DINO models. *European Journal of Mathematics and Statistics*, 4(4), 7-13. <https://doi.org/10.24018/ejmath.2023.4.4.230>
- Wu, D., Yang, J., Wan, Z. H., Shen, Y., Liu, Q., Zhang, J., ... & Li, H. (2025). A behavioral and fNIRS comparative study of gender and task differences in mental rotation among primary students. *Brain and Behavior*, 15(3), e70358. <https://doi.org/10.1002/brb3.70358>
- XUE, X., CUI, W., & ZHANG, L. (2025). A three-level meta-analysis of gender differences in spatial navigation ability. *Advances in Psychological Science*, 33(5), 843. <https://doi.org/10.3724/sp.j.1042.2025.0843>
- Yekta, M. R., & Shafiabadi, A. (2021). Qualitative analysis of factors affecting internet addiction of 15–30-year-old female victims of Iran's cyber police among ethical cases in the 2021 year. *Social Order Quarterly*, 13(3), 163–192. DOI: 20.1001.1.20086024.1400.13.3.6.4
- Zand-Moghadam, A., & Khanlarzadeh, N. (2020). Exploring the construct of academic literacy as perceived by Iranian EAP teachers. *Iranian Journal of English for Academic Purposes*, 9(3), 73-90. https://ijeap.cmu.ac.ir/article_114741.html?lang=en