Proceedings of the 90th Annual International Meeting of the Psychometric Society, Minneapolis, United States (2025).

Demographic Influences on Spatial Ability: A Rasch Model Approach

Justice Dadzie [*†], Daniel O. Oyeniran [†], Joni M. Lakin [†]

†University of Alabama, Tuscaloosa, 35401, Alabama, USA
*Corresponding author. Email: jdadzie@crimson.ua.edu

Abstract
This study applied Rasch modeling to evaluate the psychometric properties of three spatial assessments Mental Rotation Test (MRT), Object Assembly (OA), and Surface Development (SD) and to examine potential demographic differences and differential item functioning (DIF). Participants were 123 students from Grades 3–8 enrolled in a university-affiliated enrichment program. Item-level analyses indicated acceptable model–data fit (infit MSQ ≈ 0.70–1.30), with Wright maps showing good targeting of items to mid-range ability levels. Gender and race comparisons based on Rasch/EAP theta scores revealed largely overlapping ability distributions and small, nonsignificant differences across tasks. However, a Rasch tree analysis detected significant DIF by gender ($p < .001$) and by race within males ($p = .002$). Directional inspection of item parameters showed that several items were relatively easier for females on MRT and OA, whereas a subset favored males on SD. Within males, items showing race-based DIF tended to favor Non-White participants. These findings highlight the value of combining latent-scale analyses with recursive partitioning to detect and interpret subgroup-specific item bias, supporting fairer measurement of spatial ability across diverse populations.

Keywords: Spatial Ability, Rasch Model, Differential Item Functioning, STEM Education

Spatial ability is a vital cognitive skill linked to problem-solving, visualization, and reasoning across STEM fields, architecture, and design (Buckley et al., 2018; Andersen, 2014; Bodzin, 2011; Marunić & Glažar, 2014; Shi, Wind, & Lakin, 2023). It enables individuals to mentally rotate, visualize, and manipulate objects in cognitive skillse strongly associated with achievement in mathematics and scientific reasoning (Cheng & Mix, 2014; Xie et al., 2020). This study investigates how spatial ability varies across demographic groups using three assessments: the Mental Rotation Test (MRT) (Shepard & Metzler, 1971; Voyer et al., 1995), Object Assembly (OA) (Mitolo et al., 2015; Kell & Lubinski, 2013), and Surface Development (SD) (Buckley et al., 2018).

These instruments measure both dynamic and static spatial processes (Dehn, 2022; Mix et al., 2016). Building on previous findings that highlight individual differences in spatial strategy and underachievement among spatially gifted students (Hegarty, 2018; Gohm et al., 1998; Voyer et al., 2020), the present research integrates these assessments within a Rasch modeling framework to enhance measurement precision and fairness. Rasch analysis was used to estimate item difficulty and person ability for the MRT and OA, while Rasch tree analysis identified subgroup differences and differential item functioning (DIF) across gender and race (Strobl et al., 2015; Effatpanah et al., 2025; Rajeb et al., 2024). Grounded in literature connecting spatial ability to academic potential and STEM engagement (Atit et al., 2022; Froiland & Davison, 2020; Lakin & Wai, 2024), this study addresses two key questions: (1) To what extent do these items fit a Rasch model? (2) Are there mean differences in demographic groups? and (3) Do the MRT, OA, and SD exhibit DIF across gender and racial groups in cognitive and spatial processing tasks?

2. Methods

2. 1 Rasch Analysis

A Rasch modeling approach was used to estimate item difficulty, person ability, and model fit for assessing spatial ability and demographic effects. The model predicts the probability that a person with ability $\theta_n$correctly answers an item of difficulty β_i:

$$P(X_{ni} = 1) = \frac{e^{(\theta_n - \beta_i)}}{1 + e^{(\theta_n - \beta_i)}}$$

Parameters were estimated using conditional maximum likelihood, and model-data fit was assessed with infit, and outfit mean square (MSQ) statistics (acceptable range: 0.7–1.3). Items outside this range were reviewed for validity.

To detect subgroup differences, we used **Rasch tree analysis,** which provides an exploratory approach to detecting differential item functioning (DIF) along intersectional demographic groups (e.g., race within gender groups; Krist et al., 2025). Rasch tree analysis integrates recursive partitioning with the Rasch model (Strobl, Kopf, & Zeileis, 2015):

$$P(X_{ni} = 1 \mid \theta_n, \beta_i, Z_n) = \frac{e^{\left(\theta_{n,g} - \beta_{i,g}\right)}}{1 + e^{\left(\theta_{n,g} - \beta_{i,g}\right)}}, g = 1,2, \dots, G$$

Here, $Z_n$represents covariates (e.g., gender, race), and each subgroup $g$has its own Rasch model. Likelihood ratio tests ($p < .05$) identified significant splits, revealing subgroup-specific item functioning and supporting fairer spatial ability measurement.

2. 2 Participants

A total of 123 students in grades 3–8 participated in this study, recruited from university-affiliated summer camp programs in the Southeastern and Midwestern United States. These camps offered diverse educational experiences, providing a broad participant pool. Parents or guardians registered their children via email and supplied demographic data, including age, gender, and race. Participants received a gift card incentive. The final sample comprised 75 males (61%) and 48 females (39%), with 73% White (n = 90), 12% Asian (n = 15), 6.5% Native American (n = 8), 3.3% Black or African American (n = 4), 1.6% Hispanic (n = 2), and 3.3% Other (n = 4).

Three spatial ability assessments were used: the Mental Rotation Test (MRT) (17 items; Shepard & Metzler, 1971; Vandenberg & Kuse, 1978), Object Assembly (OA) (15 tasks), and Surface Development (SD) (17 items). The MRT measured dynamic spatial visualization, the OA assessed spatial organization and part–whole integration, and the SD evaluated 2D-to-3D transformations. Item complexity was systematically varied by rotation angle, occlusion, and figure orientation following Eliot and Smith (1983). All tasks were pilot-tested for clarity and reliability.

2. 3 Procedures

Assessments were administered in classroom settings under standardized conditions, with sessions lasting about 30 minutes. Rasch modeling was applied to the MRT and OA to estimate item difficulty and person ability, while the SD was analyzed separately. Rasch tree analysis, t-tests, and ANOVAs examined demographic effects, ensuring fair, reliable, and valid measurement of spatial reasoning across gender and racial groups (Bond & Fox, 2015; Strobl et al., 2015; Xie et al., 2020).

3. Results

The results are organized in two parts. The first part presents the Rasch model findings related to item difficulty and model fit for the Mental Rotation Test (MRT), Object Assembly (OA), and Surface

Development Test (SD). The second part reports demographic analyses, including gender and racial group differences across the MRT, OA, and Surface Development (SD) tasks.

3. 1 Item Difficulty and Model Fit

Across the three item-fit plots, most items showed acceptable model–data fit, clustering near the expected infit MSQ of 1.0 and remaining within the dashed acceptable band (≈0.70–1.30). For the MRT (Figure 1), item difficulties ranged from very easy to moderately hard, and fit was generally stable; however, **mrt6** and **mrt16** exceeded the upper infit boundary, indicating underfit, while **mrt4** fell below the lower boundary, suggesting possible overfit. For the OA (Figure 2), difficulties spanned a wider range, and most items fit well, but **oa13** and **oa7** approached or exceeded the upper boundary, and **oa2** was near the lower boundary. In the Surface Development plot (Figure 3), most items fit adequately, with a small number (e.g., **T1_Bscore**) showing notable underfit. Wright maps confirmed strong item–ability alignment, establishing all tests as reliable, complementary measures of spatial reasoning (Bond & Fox, 2015; Buckley et al., 2018; Rajeb et al., 2024).

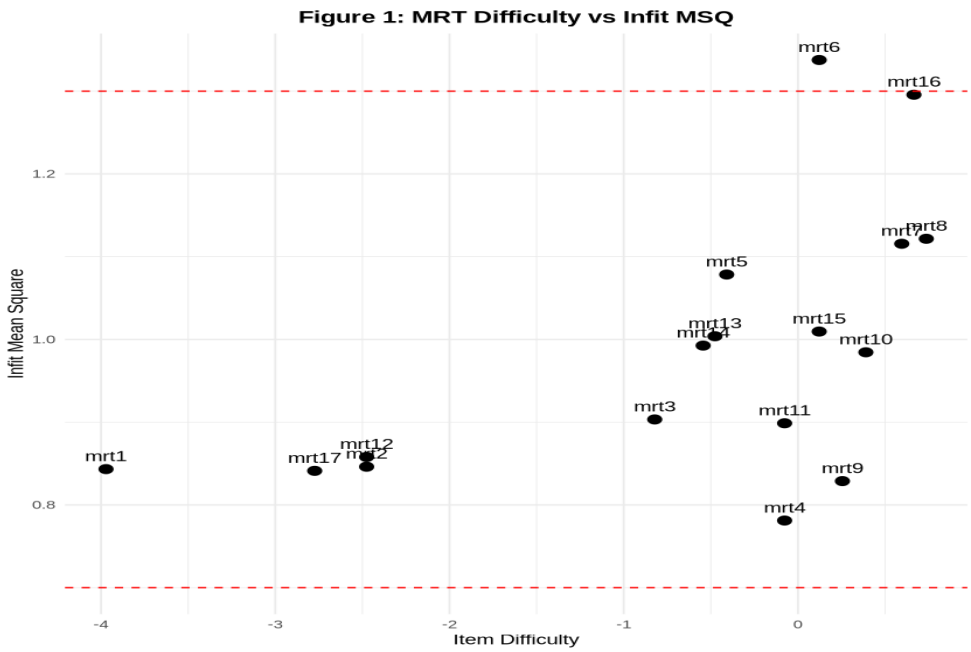

Figure 1. Item Fit Plot: Relationship Between Item Difficulty and Infit Statistics for the Mental Rotation Test (MRT)

Note. Each point represents an MRT item. Items within the 0.7–1.3 infit range indicate acceptable fit; deviations suggest potential multidimensionality or misfit.

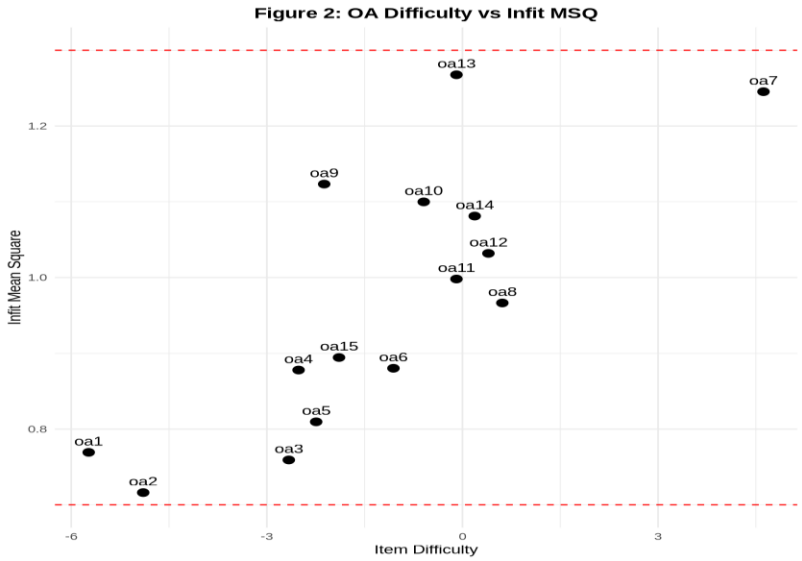

Figure 2. Item Fit Plot: Relationship Between Item Difficulty and Infit Statistics for the Object Assembly (OA)

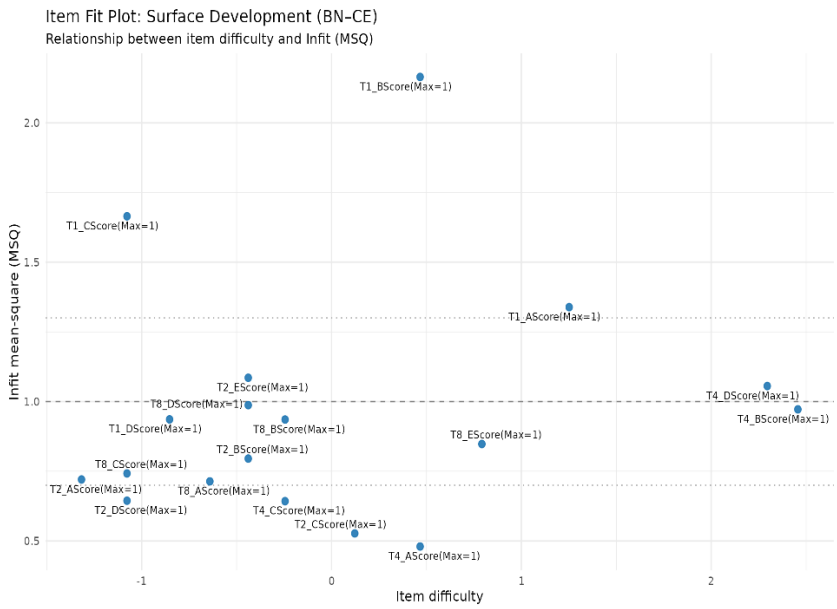

Figure 3. Item Fit Plot: Relationship Between Item Difficulty and Infit Statistics for the surface development (SD)

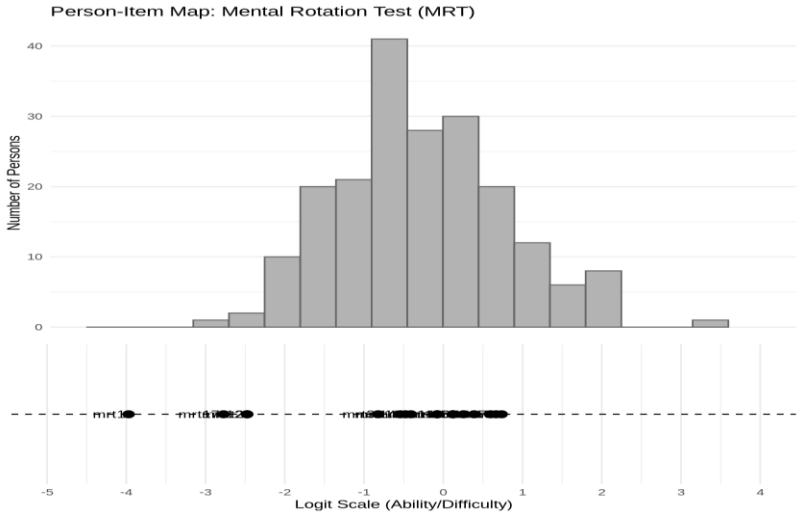

Figure 4. Wright Map for the Mental Rotation Test (MRT)

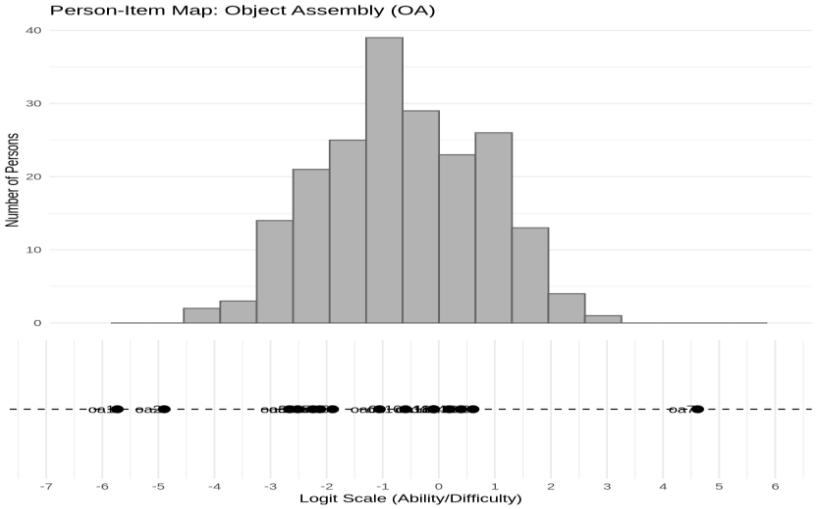

Figure 5. Wright Map for the Object Assembly (OA)

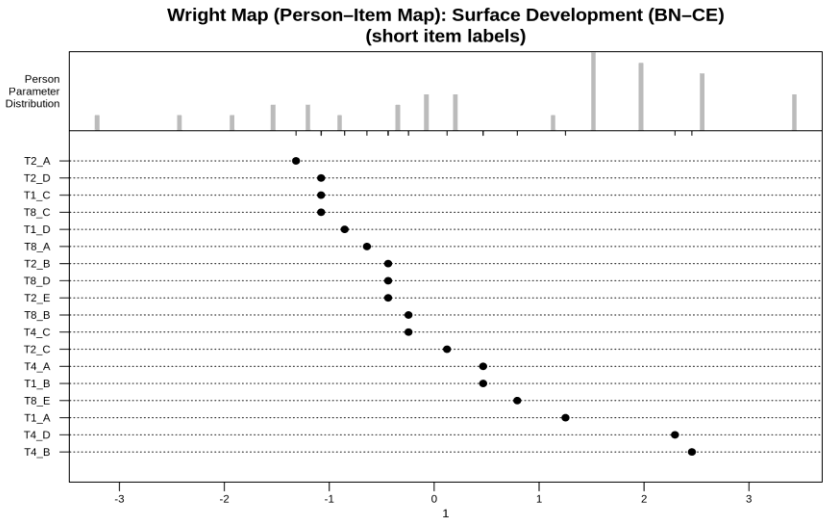

Figure 6. Wright Map for the surface development (SD)

3. 2 Demographic Differences in Spatial Ability

Figure 7 shows substantial overlap in Rasch/EAP theta distributions for females and males on MRT and OA, with little evidence of gender separation. This pattern matches the non-significant tests for MRT, $t(74.23) = -0.283$, $p = .778$, $d = -0.064$, and OA, $t(59.14) = -0.964$, $p = .339$, $d = -0.223$. SD shows a clearer rightward shift for males, although still overlapping and non-significant, $t(50.60) = -1.538$, $p = .130$, $d = -0.416$. Figure 8 suggests small-to-moderate shifts favoring Non-White participants on MRT and OA (marginal *ps*), but near-complete overlap on SD (Hegarty, 2018; Voyer et al., 2020).

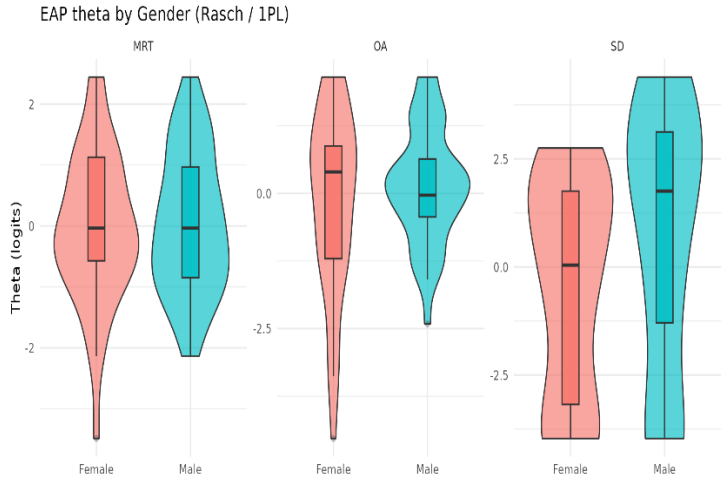

Figure 7: Distributions of Rasch (1PL) EAP Theta Scores by Gender Across MRT, OA, and SD

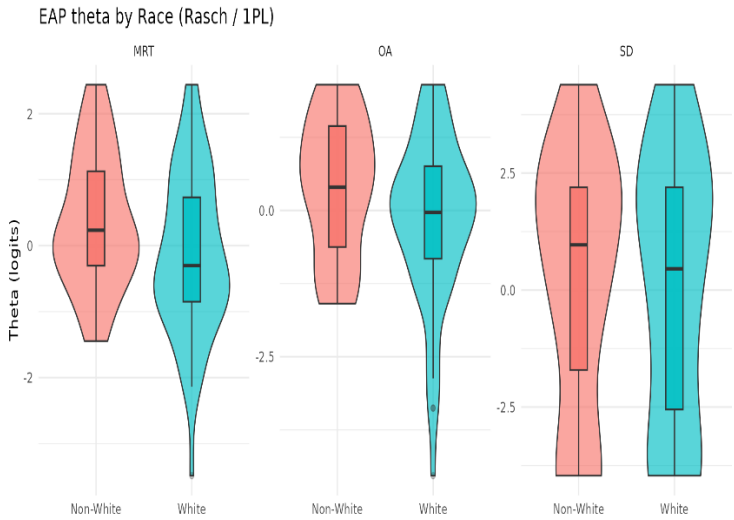

Figure 8: Distributions of Rasch (1PL) EAP Theta Scores by Race Across MRT, OA, and SD

3.2.1 Rasch Tree Analysis

The Rasch tree analysis revealed significant DIF across demographic groups. The first split by gender ($p < .001$) indicating noninvariance in item parameters between females and males even after conditioning on the latent trait. Examination of group-specific item difficulty estimates showed that most of the items contributing to this split were easier for females on the MRT and OA, suggesting that these items favored female respondents. In contrast, a small subset of SD items was easier for males, indicating localized DIF favoring males on that task. Within the male subgroup, a second split by race ($p = .002$) revealed

additional DIF, with the affected items generally easier for Non-White males, suggesting slight item-level advantages for this group. (Strobl et al., 2015; Effatpanah et al., 2025).

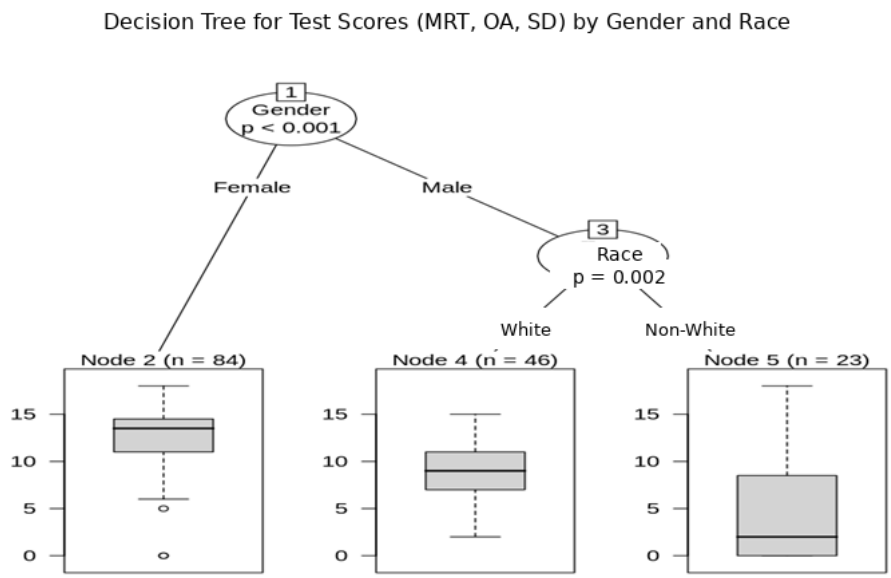

Figure 5. Rasch Tree Analysis of Demographic Subgroup Differences in Spatial Ability
Note. The Rasch tree displays hierarchical splits first by gender and then by race within the male subgroup, indicating subgroup-specific variation in item functioning and spatial performance.

4. Discussion

This study examined spatial ability using a Rasch modeling framework to explore item-level difficulty patterns and demographic influences across three spatial assessments: the Mental Rotation Test (MRT), Object Assembly (OA), and Surface Development (SD). The findings demonstrated that the Rasch model provided a robust analytic framework for evaluating item performance, identifying subgroup differences, and ensuring measurement fairness. Consistent with prior work (Bond & Fox, 2015; Boone, Staver, & Yale, 2014), the Rasch model allowed for meaningful interpretation of item difficulty and person ability on a common logit scale, while the Rasch tree analysis extended this understanding by identifying differential item functioning (DIF) across gender and racial subgroups.

4. 1 Item Difficulty and Model Fit

The Rasch analysis revealed that both the MRT and OA demonstrated acceptable psychometric quality, with most items fitting the model expectations. A small number of items in each test showed misfit, suggesting multidimensionality or unexpected response patterns. These findings align with previous studies reporting that spatial ability measures often capture overlapping but distinct cognitive processes such as rotation, visualization, and spatial structuring (Buckley, Seery, & Canty, 2018; Uttal et al., 2024). The OA showed a wider difficulty range than the MRT, implying that the task captured more diverse cognitive demands, including pattern construction and mental synthesis. Similar observations have been made by Carroll (1993) and Linn and Petersen (1985), who noted that spatial reasoning encompasses multiple subskills that may not load onto a single dimension. The Wright maps confirmed adequate targeting between item difficulty and participant ability levels, demonstrating that both instruments effectively measured a broad spectrum of spatial skills.

4. 2 Gender and Race Difference in Spatial Ability

Gender- and race-related patterns in spatial ability appeared more nuanced than a uniform subgroup-gap narrative. On the Rasch/EAP latent scale, MRT and OA showed substantial overlap between females and males, offering little evidence of meaningful separation. This pattern challenges claims of a consistent male advantage (Voyer, Voyer, & Bryden, 1995) and aligns with work emphasizing that gender differences are malleable and context-dependent across tasks and learning histories (Hegarty, 2018; Voyer et al., 2020; Krist et al., 2025). The limited separation is also consistent with evidence that spatial skills are trainable and responsive to instruction and strategy development (Nazareth, Herrera, & Pruden, 2019; Uttal et al., 2013). Racial variability was similarly task-specific: White versus Non-White distributions suggested small-to-moderate shifts on MRT and OA, but near-complete overlap on SD, echoing research linking spatial outcomes to opportunity and enriched learning environments (Kell & Lubinski, 2013; Yang et al., 2020; Sorby et al., 2022; Xie et al., 2020). Notably, the Rasch tree indicated potential item-level noninvariance, with gender as the primary split and race within males, underscoring the value of DIF-sensitive diagnostics for fair interpretation (Strobl, Kopf, & Zeileis, 2015; Effatpanah, Ravand, & Doebler, 2025).

5. Conclusion

Overall, this study provides empirical evidence that the Rasch modeling approach offers a precise and equitable framework for assessing spatial ability while accounting for demographic variability. The combination of Rasch and Rasch tree analyses enabled a deeper understanding of item functioning and subgroup performance differences. The results emphasize that spatial ability is not a fixed trait but a multifaceted, experience-sensitive skill shaped by educational opportunities, cultural exposure, and gendered learning contexts. These insights support the ongoing development of fair, valid, and inclusive spatial assessments that better reflect the diverse strengths of learners across demographic backgrounds.

6. Implications for Assessment and Education

The findings hold important implications for both spatial assessment and STEM education. First, the results affirm that Rasch modeling can serve as a powerful diagnostic tool for evaluating item quality and participant performance. By identifying misfitting items, educators and test developers can refine spatial measures to achieve more valid and equitable assessments. Second, the observed demographic effects highlight the need for culturally responsive and gender-inclusive approaches in spatial education. As demonstrated by Sorby (2009) and Uttal et al. (2013), targeted spatial training interventions can significantly enhance performance, particularly among groups historically underrepresented in spatially intensive domains. Incorporating such interventions into STEM curricula may help reduce performance disparities and foster equitable development of spatial reasoning skills.

7. Limitations and Future Directions

Several limitations should be noted. The sample size, though sufficient for Rasch estimation, was relatively small and drawn from university-affiliated summer programs, which may limit generalizability. Future studies should include larger, more diverse populations and longitudinal designs to examine how demographic effects evolve over time. Additionally, the study relied on cross-sectional data; thus, causal interpretations regarding educational exposure and spatial development remain speculative. Further research should also explore the multidimensional structure of spatial ability through multidimensional Rasch modeling to capture interrelated but distinct cognitive components such as mental rotation, object manipulation, and surface visualization.

Acknowledgement

The authors would like to express their deepest gratitude to Dr. Stefanie Wind for her invaluable support, guidance, and encouragement throughout the development of this manuscript and its presentation at the conference. We also extend our sincere appreciation to all our colleagues for their continual support, collaboration, and constructive suggestions during the research and writing process.

References

Andersen, L. (2014). Visual–spatial ability: Important in STEM, ignored in gifted education. *Roeper Review*, *36*(2), 114-121. https://doi.org/10.1080/02783193.2014.884198

Atit, K., Power, J. R., Pigott, T., Lee, J., Geer, E. A., Uttal, D. H., ... & Sorby, S. A. (2022). Examining the relations between spatial skills and mathematical performance: A meta-analysis. *Psychonomic bulletin & review*, 1-22. https://doi.org/10.3758/s13423-021-02012-w

Bodzin, A. M. (2011). The implementation of a geospatial information technology (GIT)-supported land use change curriculum with urban middle school learners to promote spatial thinking. *Journal of Research in Science Teaching*, *48*(3), 281-300. **https://doi.org/10.1002/tea.20409**

Buckley, J., Seery, N., & Canty, D. (2018). A heuristic framework of spatial ability: A review and synthesis of spatial factor literature to support its translation into STEM education. *Educational Psychology Review*, *30*(3), 947-972. https://doi.org/10.1007/s10648-018-9432-z

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies* (No. 1). Cambridge university press.

Cheng, Y. L., & Mix, K. S. (2014). Spatial training improves children's mathematics ability. *Journal of cognition and development*, *15*(1), 2-11. https://doi.org/10.1080/15248372.2012.725186

Dehn, M. J. (2022). *Essentials of processing assessment*. John Wiley & Sons.

Effatpanah, F., Ravand, H., & Doebler, P. (2025). Differential Item Functioning Analysis of Likert Scales: An Overview and Demonstration of Rating Scale Tree Model. *Psychological Reports*, 00332941241308806. https://doi.org/10.1177/00332941241308806

Froiland, J. M., & Davison, M. L. (2020). Social perception: relationships with general intelligence, working memory, processing speed, visual-spatial ability, and verbal comprehension. *Educational Psychology*, *40*(6), 750-766. https://doi.org/10.1080/01443410.2020.1732873

Gohm, C. L., Humphreys, L. G., & Yao, G. (1998). Underachievement among spatially gifted students. *American Educational Research Journal*, *35*(3), 515-531. https://doi.org/10.3102/00028312035003515

Hawes, Z. C., Gilligan-Lee, K. A., & Mix, K. S. (2022). Effects of spatial training on mathematics performance: A meta-analysis. *Developmental Psychology*, *58*(1), 112. https://doi.org/10.1037/dev0001281

Hegarty, M. (2018). Ability and sex differences in spatial thinking: What does the mental rotation test really measure?. *Psychonomic bulletin & review*, *25*, 1212-1219. https://doi.org/10.3758/s13423-017-1347-z

Hegarty, M., & Waller, D. (2005). Individual differences in spatial abilities. *The Cambridge handbook of visuospatial thinking*, 121-169.

Heo, M., & Toomey, N. (2020). Learning with multimedia: The effects of gender, type of multimedia learning resources, and spatial ability. *Computers & Education*, *146*, 103747. https://doi.org/10.1016/j.compedu.2019.103747

Holden, M. P., Newcombe, N. S., & Shipley, T. F. (2015). Categorical biases in spatial memory: the role of certainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 473.

Kell, H. J., & Lubinski, D. (2013). Spatial ability: A neglected talent in educational and occupational settings. *Roeper Review*, *35*(4), 219-230. https://doi.org/10.1080/02783193.2013.829896

Krist, A. T., Wind, S. A., & Lakin, J. M. (2025). Rasch tree modeling for DIF analysis: how sample and item characteristics influence DIF detection. *Applied Measurement in Education*, *38*(2), 149-164. https://doi.org/10.1080/08957347.2025.2559811

Lakin, J. M., & Wai, J. (2020). Spatially gifted, academically inconvenienced: Spatially talented students experience less academic engagement and more behavioural issues than other talented students. *British Journal of Educational Psychology*, *90*(4), 1015-1038. **https://doi.org/10.1111/bjep.12343**

Marunić, G., & Glažar, V. (2014). Improvement and assessment of spatial ability in engineering education. *Engineering Review: Međunarodni časopis namijenjen publiciranju originalnih istraživanja s aspekta analize konstrukcija, materijala i novih tehnologija u području strojarstva, brodogradnje, temeljnih tehničkih znanosti, elektrotehnike, računarstva i građevinarstva*, *34*(2), 139-150.

Mitolo, M., Gardini, S., Caffarra, P., Ronconi, L., Venneri, A., & Pazzaglia, F. (2015). Relationship between spatial ability, visuospatial working memory and self-assessed spatial orientation ability: a study in older adults. *Cognitive processing*, *16*, 165-176. https://doi.org/10.1007/s10339-015-0647-3

Mix, K. S., Levine, S. C., Cheng, Y. L., Young, C., Hambrick, D. Z., Ping, R., & Konstantopoulos, S. (2016). Separate but correlated: The latent structure of space and mathematics across development. *Journal of Experimental Psychology: General*, *145*(9), 1206. https://doi.org/10.1037/xge0000182

Rajeb, M., Krist, A. T., Shi, Q., Oyeniran, D. O., Wind, S. A., & Lakin, J. M. (2024). Mental Rotation Performance: Contribution of Item Features to Difficulties and Functional Adaptation. *Journal of Intelligence*, *13*(1), 2. **https://doi.org/10.3390/jintelligence13010002**

Ramful, A., Lowrie, T., & Logan, T. (2017). Measurement of spatial ability: Construction and validation of the spatial reasoning instrument for middle school students. *Journal of Psychoeducational Assessment*, *35*(7), 709-727. https://doi.org/10.1177/0734282916659207

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701-703. https://doi.org/10.1126/science.171.3972.701

Shi, Q., Wind, S. A., & Lakin, J. M. (2023). Exploring the influence of item characteristics in a spatial reasoning task. *Journal of Intelligence*, *11*(8), 152. **https://doi.org/10.3390/jintelligence11080152**

Sorby, S. A. (2009). Educational research in developing 3-D spatial skills for engineering students. *International Journal of Science Education*, *31*(3), 459-480. https://doi.org/10.1080/09500690802595839

Sorby, S., Duffy, G., Soni, D., & Panther, G. (2022, August). An Evaluation of the Relationship between Spatial Skills and Creating a Free Body Diagram. In *2022 ASEE Annual Conference & Exposition*.

Strobl, C., Kopf, J., & Zeileis, A. (2015). Rasch trees: A new method for detecting differential item functioning in the Rasch model. *Psychometrika*, *80*(2), 289-316. https://doi.org/10.1007/s11336-013-9388-3

Uttal, D. H., McKee, K., Simms, N., Hegarty, M., & Newcombe, N. S. (2024). How can we best assess spatial skills? Practical and conceptual challenges. *Journal of Intelligence*, *12*(1), 8. **https://doi.org/10.3390/jintelligence12010008**

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

Voyer, D., Saint-Aubin, J., Altman, K., & Doyle, R. A. (2020). Sex differences in tests of mental rotation: Direct manipulation of strategies with eye-tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(9), 871. https://doi.org/10.1037/xhp0000752

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological bulletin*, *117*(2), 250. https://doi.org/10.1037/0033-2909.117.2.250

Wai, J., & Lakin, J. M. (2024). The Importance of Identifying Spatial Skills in Talented Students. *Gifted Child Today*, *47*(3), 178-185. https://doi.org/10.1177/10762175241242496

Xie, F., Zhang, L., Chen, X., & Xin, Z. (2020). Is spatial ability related to mathematical ability: A meta-analysis. *Educational Psychology Review*, *32*, 113-155. https://doi.org/10.1007/s10648-019-09496-y

Yang, W., Liu, H., Chen, N., Xu, P., & Lin, X. (2020). Is early spatial skills training effective? A meta-analysis. *Frontiers in psychology*, *11*, 1938. https://doi.org/10.3389/fpsyg.2020.01938