



Kumiko Shiina

What Spatial Tests Measure:

A Review of Validity Evidence — Methods, Findings, and Practical Directions

This paper synthesizes empirical evidence on the validity of common spatial tests, particularly the Mental Cutting Test (MCT) and the Mental Rotations Test (MRT), and outlines practical directions for translating validity research into diagnostic practice. Three complementary strands are reviewed: (1) response-process studies (eye-tracking, verbal protocols, and software logs) that reveal heterogeneous solving routes not captured by total scores, (2) relations-to-other-variables analyses that evaluate how scores relate to instructional outcomes and emphasize the necessity of appropriate comparison groups for attributing pre/post changes to instruction, and (3) internal-structure modeling (cognitive diagnostic models, Q-matrices, and G-DINA) that identifies multiple underlying spatial attributes. Integrating these strands, the paper recommends routine, low-burden logging of process traces, cross-context evaluation of candidate features, controlled study designs, and iterative development of scoring and reporting algorithms with pre-specified success criteria. Realizing these steps requires interdisciplinary collaboration, sensitivity analyses and replication, and careful attention to ethical and data-privacy issues. Concrete recommendations for logging, evaluation, and iterative validation are provided to bridge research and classroom practice.