What are learning outcomes for spatial thinking curricula: What form should assessment take?

BETH CASEY
Professor Emeritus
School of Education
Boston College
Email: beth.casey@bc.edu

The purpose of this position paper is to focus on the question of what form assessments should take when evaluating the relation between spatial thinking and STEM curricula. The first recommendation is that we need to consider effective ways of conducting evaluations of spatial interventions that do not depend solely on the random assignment of students to experimental and control conditions. I will discuss an alternative research and statistical design, the Regression Discontinuity approach. Secondly, research has now established that spatial abilities are related to success in STEM disciplines. However, we need a much more refined analysis of: (1) which specific types of content within each STEM field are impacted by spatial skills, (2) the specific mechanisms underlying these associations, and (3) a better understanding of individual differences in the spatial strategies (successful or unsuccessful) that students use when trying to solve spatially related STEM content.

An Alternative Approach to Design of Spatial Interventions at the College Level: The Regression Discontinuity Approach

The organizers of this conference make a very important point, when they state that, “…spatial thinking is not fostered in our educational system and that current practice depends more on selection of the most able students for spatially demanding disciplines than on fostering the spatial intelligence of all students.” In order to determine the effectiveness of a particular spatial intervention on college student learning, it is clearly critical to use appropriate research design methods. However, for a wide variety of reasons, not all researchers and curriculum developers can apply the gold standard of research design by randomly assigning college students to experimental and control groups.

In a recent paper that I co-authored with Sheryl Sorby and others (Sorby, Casey, Veurink, & Dulaney, 2012; under review, Learning and Individual Differences), we used a different research design to determine the effectiveness of spatial interventions on Michigan Tech engineering students at risk of poor spatial skills. The major contribution of this new methodology was to confirm and validate 15 years of prior findings on spatial skill interventions conducted at the same Engineering Program—by applying a more statistically sophisticated approach—Regression Discontinuity (RD). A problem interpreting the past results is that there might have been a selection bias, since students who failed the initial spatial test could decide to either take the intervention or serve as the comparison group. Thus, the students choosing to take the intervention may have had higher motivation levels than the students not taking it.
Our new study was designed to examine the benefits of an intervention targeted to the freshmen engineering students who failed an initial spatial assessment during orientation at Michigan Tech. It was not possible to randomly assign students to conditions. Instead, we required all students who failed the mental rotation test during orientation to enroll in the spatial intervention course. This enabled us to address the question of whether the spatial intervention was successful in raising students’ spatial skills through an alternative non-experimental design. Using a RD pre/post-test analysis, we found a treatment effect by demonstrating a discontinuity or jump in the regression intercepts at the cutoff score of the pretest variable, with the intervention group performing at higher levels than would be expected if there had been no intervention. Using the same RD analysis, the intervention also showed transfer effects, improving calculus performance of the students in the intervention condition.

One strong argument for the regression discontinuity (RD) design is that it allows for elimination of selection bias when implemented properly. In fact, in cases where a randomized design is not possible, a RD design is the recommended alternative to quasi-experimental and associational designs because it allows for an unbiased detection of treatment effects (Cook, 2008; Institute of Educational Sciences (IES), Technical Methods Report, 2008; Shadish, Galindo, Wong, Steiner, & Cook, 2011). The RD design is based on a pretest-posttest treatment-comparison group design, in which individuals are assigned to a treatment condition based on a cutoff score from a pre-intervention measure. Participants scoring on one side of the cutoff receive the intervention while participants scoring on the other side of the cutoff do not receive the intervention. As long as assignment to the intervention and comparison conditions strictly follows the cutoff criterion, any selection effects correlated with the impact of the intervention are also perfectly correlated with the pre-intervention measure, which, when held constant in the statistical analysis, allows for an unbiased estimate of the intervention impact on a post-intervention measure (Shadish, Cook, & Campbell, 2001; Thistlethwaite & Campbell, 1960; Trochim & Cappelleri, 1992). Consequently, researchers are starting to use the RD design to obtain unbiased impact estimates of education-related interventions when random assignment is not possible (IES, 2008). Like the experimental design, the logic underlying the RD design supports statistically valid conclusions, as evidenced by statistical proofs (Cappelleri, 1991; Rubin, 1977). In addition to eliminating selection bias, the RD design also avoids problems of regression to the mean presented by a cross-sectional analysis. Therefore, without selection effects and regression to the mean, the RD design avoids threats to internal validity that are inevitably posed by the cross-sectional design, making it a useful approach. (For a complete discussion of the RD design and internal validity, see Shadish et al., 2001).

A More Refined/In-Depth Approach to Understanding the Relation between Spatial Skills and STEM Performance

To make further progress in understanding the spatial/STEM content relationship, it is important to move beyond establishing an association between performance on a spatial measure and global achievement measures for different STEM fields. In my view, this more fine-grained analysis is an important first step that should be conducted prior to the design of spatial
interventions within specific disciplines. For example, in a recently funded NSF proposal on the relation between spatial skills and middle school students’ math achievement, we plan to address the question of whether a spatial/math association generalizes to math content as a whole or only to specific math content. Thus, we plan to investigate the relation between spatial skills and different types of mathematics achievement: (a) content likely to depend more on analytical, logical-deductive reasoning, and (b) content likely to depend more on spatial reasoning. This type of fine-grained approach may help to provide a clearer specification of the mechanisms by which spatial training could lead to improvement in different content areas within the specific STEM disciplines.

Secondly, when examining spatial skills, we need to develop more effective methods for identifying the strategies that students use when approaching spatial problem solving. For example, Geiser and his associates (Geiser, Lehmann, & Eid, 2006) developed a useful method for identifying individual differences in spatial strategy use by applying latent class analysis to identify groups of students whose response patterns were highly similar (making it possible to identify the strategies they used for solving the spatial problems). Then we need to use these methodologies to determine whether individual differences in spatial strategy use can be linked to variations in levels of performance on different types of content, concepts, and problem solving approaches within a discipline. Ultimately, it is important to determine whether changes in spatial strategy choice (e.g., moving from more analytical to more holistic approaches) are connected to changes and improvement in students’ performance following spatial interventions.