In an era in which information is rapidly growing and changing, it is very important to teach with the goal of students' engagement in life-long learning in mind. This can partially be achieved by developing transferable thinking skills. Transfer refers to students' ability to apply knowledge and skills in new learning contexts. We formulated a theoretical transfer framework that distinguishes between near and far transfer. The framework consists of three attributes: task distance, interdisciplinarity, and skills set. The goal of this research was to explore the applicability of the three-attribute transfer skills framework in the design of the transfer assignments as a tool to assess the development in students' skills. About 670 chemistry 12th grade honor students from 24 high schools in Israel participated in the research. The students learned a special program in a computerized laboratory setting. Findings indicated an increase in students' far transfer skill as expressed by the progress students made in transferring knowledge from chemistry to other science domains and by using more chemistry understanding levels in their responses. The research demonstrates the use of the theoretical framework as a tool to mediate between theory and practice, the framework provides a unique interface between researchers and teachers in science education. Teachers may become ‘experts’ by serving as designers of learning environments by producing their own transfer assignments. They can use this method to conduct ‘action research’ to investigate the relationships between their pedagogical choices and their students’ learning outcomes.

Validation of a large-scale test measuring pre-service science teachers' scientific reasoning skills

Psychometrics,Quantitative methods,Assessment methods and tools,Pre-service teacher education,Competencies,Science education

Stefan Hartmann, Humboldt-Universitat zu Berlin , Germany;

The aim of this study was to initially validate a multiple-choice test which was developed to evaluate the development of pre-service science teachers' scientific reasoning skills. Three hypotheses regarding differences between extreme groups were tested. A sample of 2156 academic students in teacher education programs participated on the study. The test contained 123 multiple-choice items that were assigned to test booklets using a balanced incomplete block design. A two-parametric IRT model was used to estimate the students' abilities. To minimize the impact of possible confounding effects on the results, propensity score matching was applied to the sample prior to the analysis of group differences. As predicted, students at the end of academic training performed higher than students at the beginning, students who study two natural sciences performed higher than students who study only one natural science, and students who attended seminars in which scientific reasoning is taught explicitly performed higher than students who did not attend such seminars. The evidence found in this study provides support for the criterion-based validity of our interpretation of the test scores as measures of scientific reasoning competencies.

Online Assessment of Combinatorial Reasoning: Perspectives of Measuring a Challenging Construct

Quantitative methods,Assessment methods and tools,Educational technology,Achievement,Cognitive skills,Reasoning

Beno Csapo, University of Szeged, Hungary; Attila Pasztor, University of Szeged, Hungary; Gyongyver Molnar, University of Szeged, Hungary;

This paper presents results from the development of an online test for assessing combinatorial reasoning using innovative item formats and automated scoring. The objectives of the study are to analyze the psychometric characteristics of the online instrument, examine age and gender differences, and explore the possibilities of further analyses. The sample was drawn from students in third (N=186) and fourth (N=219) grades. An online test was developed based on a former paper-and-pencil test that measured a number of operations of combinatorial reasoning with two types of content (formal content with letters and numbers and figural content with pictures). Online data collection was carried out by means of the eDia (Electronic Diagnostic Assessment) platform. Students entered their responses by keyboard (formal content) or by dragging and dropping figures on the screen with a mouse (figural content). The reliability indices for the test were good (Cronbach's ? = .84 for figural and .88 for figural and formal items). Mplus was used to analyze model fit and scale invariance; the results provided evidence of the test's construct validity. Grade and gender differences were found: girls achieved significantly better test scores than boys, and fourth graders performed better than third-grade students. The differences between the third and fourth graders indicate that there may be structural differences between younger and older children's thinking skills. The findings indicate that online assessment may provide teachers with an easy-to-use instrument for monitoring the development of students' combinatorial reasoning.

Computer-based assessment of creativity: The case of divergent thinking

Qualitative methods,Assessment methods and tools,Educational technology,Achievement,Cognitive skills,Reasoning

Attila Pasztor, University of Szeged, Hungary; Gyongyver Molnar, University of Szeged, Hungary; Beno Csapo, University of Szeged, Hungary;

The purpose of this study is to explore the possibilities of technology-based assessment of divergent thinking and to contribute to the development of a reliable online instrument. The sample for the study was drawn from sixth-grade students (N=917). The computerized instrument comprised six tasks and was based on Torrance's and Wallach and Kogan's item types for divergent thinking. Answers were scored by the scales of fluency, flexibility and originality. The online data collection was carried out by the eDia platform via Internet in schools' ICT rooms. The comparison was made between the respondent database and the database with the categorized answers using a computer program which calculated the three indices automatically. The online assessment tool proved to be reliable (Cronbach's alpha ranged between .80 and .87). Correlation patterns between verbal and figural subtests provided empirical evidence for the convergent and discriminant validity. The three-dimensional model based on fluency, flexibility and originality fit the data better than the one-dimensional model (Chi-square=972.54, df=3 p.001), thus factors of divergent thinking distinguished in paper-and-pencil testing could be empirically distinguished in a computer-based environment as well. Within the three-dimensional model, all three dimensions correlated on a latent level (r_flu_fle=.61, r_flu_or=.65, r_fle_or=.80, p.001). Online test administration and automatic scoring reduced time