Development and empirical validation of a model of music-related argumentative competence

Background & theoretical competence model

In many countries learning outcomes have been formulated in the form of achievement standards for several school subjects. In Germany, it is required that standards be developed on the basis of competence models (cf. Klieme & Maag Merki, 2008). For music, an initial attempt has been made with regard to the competence of “perceiving and contextualising music” (Jordan & Knigge, 2010). Based on these results, our study focuses on another important competence: the aim is to develop and empirically validate a competence model of music-related argumentation. The starting point is a theoretical model (Rolle, 2013) that includes three dimensions (distinguishing the abilities to identify, produce, and evaluate arguments in disputes about music). Each dimension is divided into seven levels. The model structure refers to a developmental sequence in the way people judge artworks (Parsons, 1987) and to King and Kitchener’s (1994) “reflective judgement model”.

Phase I: Model specification

In order to differentiate and revise the model we conduct a qualitative study with a small sample of grade nine students. The combination of an experimental approach (discussion tasks in an internet forum) with focus group discussions produces data that will be analysed using argumentation and conversation analysis.

Phase II: Operationalization of the model

The operationalization of the (revised) model by means of specific tasks (= item development process) has the following circular structure:

i. Monthly meeting of an expert group (teachers and academics) for item development (and revision).
ii. Transformation of items into a technology based assessment system (ILIAS).
iii. Online piloting of the items with respect to their comprehensibility and the clarity of directions and scoring manuals, etc.; sample size: approx. 200 grade nine students of expert group teachers.
iv. Students’ approaches to solving the tasks are to be investigated in cognitive labs (“think aloud” method).
v. Quantitative (e.g. item difficulty/fit/discrimination) and qualitative analyses of the tests and cognitive labs.

Phase III: Validation Study

A multi-matrix-design is used to pilot more items than a single student is able to answer in a test (cf. Gonzalez & Rutkowski, 2010). Therefore the items will be spread over several test-booklets, all linked with anchor items. The study sample consists of grade nine students representing all types of German schools. The data will be analysed using item response theory methods (e.g. Embretson & Reise, 2000):

i. Item analysis: e.g. item fit/difficulty/discrimination, DIF.
ii. Structural Validity: model structure analyses with multidimensional IRT models and comparisons of model fit (e.g. CAIC).
iii. Proficiency level definition: setting cut-scores by using “task demands”.
iv. External validity: background variables as well as closely connected constructs (e.g. literary-aesthetic judgement competence) will be analysed using structural equation models.

(Anticipated Results)

i. Model with validated competency dimensions and levels.
ii. Assessment instrument that can subsequently be used for educational monitoring, evaluation, etc.