Discouraging results: problematizing test questions in science education

Paper presented at ESERA conference Cyprus 2-7th of September 2013
Author: Margareta Serder, Malmo University, Sweden. E-mail: margareta.serder@mah.se

Background, review of research and theoretical point of departure
Worldwide, the field of education experiences the increasingly stressed importance of test-driven practice (Broadfoot & Black, 2004) of which the consequences are many, such as accountability for teachers, schools, and states (Ravitch 2010), and narrowing practice towards “teaching for the test” and subsequently lower the level of education (Posner, 2004). The importance of testing is furthermore emphasized when international comparisons, such as the OECD product Programme International of Student Assessment (PISA), are regarded as appropriate tools for supporting educational reforms and practices (Pons, 2012). The size of the consequences calls for a need to problematize tests and test-questions, by which skills, ability, knowledge, and literacies are assessed.

In PISA, the largest current knowledge assessment, 15-year old students’ knowing of and about science is tested, framed as the students’ scientific literacy (OECD, 2006). This framework emphasizes that the knowledge is to be used in, and relevant for, general life and therefore the test questions are “situated” in written descriptions of everyday contexts. PISA is thus a worldwide promoter for science learning and science for citizenship, and meanwhile relying on an immense system of individual knowledge measurement to assess this.

Lemke (1990) argues that learning science is to learn the language of science. According to Yore and Treagust (2006) science learners are confronted with a “three-language problem”; i.e. the need of mastering and making distinction between the discourse of home, instruction, and science: different discourses are at play in different locations and contexts. However, as pointed out by Linell (1998), what the context is and consequently the appropriate discourse is not obvious.

Research has indicated that the pre-defined science competencies the PISA test claims to measure often not correspond to those actually used by the students (Lau, 2009). Others have demonstrated that the way test-questions are formulated in themselves largely affect the ways the students perceive and answer to them (Olsen et al., 2001; Schoultz et al., 2001).

This paper adopts a dialogical theoretical perspective (Linell, 1998) and aims to discuss one of the assumptions that individual assessment relies on, namely that of message-transfer (ibid). It does so through exploring the interaction between students and test-questions in scientific literacy, with PISA as an example. How are the texts, illustrations and everyday context in the test negotiated by the students, and what meanings are students’ making from them in a situation of collaborative problem-solving?

Methodology, analysis, and findings
In order to dialogically (in line with Linell, 1998) explore what meanings can be activated by scientific literacy test items, a research design was chosen in which students’ meaning making could be observed in action, that is, during the work with the questions to answer. Therefore 15 year olds, from four different Swedish 9th grade science classes, were asked to work with and respond to a selection of PISA Science items in teacher constructed,
heterogeneous, groups of 2-4 students while their talk and actions were videotaped. The data consists of 16 hours of video recordings together with the groups’ written answers. The units from the released PISA Science items 2006 (Acid Rain [S485], Greenhouse – fact or myth [S114] and Sunscreen [S447]) were selected as they met the criteria of representativeness, relevance and long term use in the PISA assessment.

The talk and actions were transcribed in Swedish and thereafter a content analysis was conducted. The transcripts were also semantically analyzed using the analytic model of Mäkitalo, Jakobsson, and Säljö (2009). This analysis focused gaps in students’ interaction (Wickman, 2004), analysing students’ negotiations of the meanings being made in different situations. Such gaps were analyzed in terms of meaning potential (Rommetveit, 1974).

The content analysis shows that all groups (more or less) negotiate the very meaning of the test questions: What is meant by certain formulations? What is the intended meaning of test-occurring words like “factor”, “reference”, “pattern”, “better” or “constant”? The semantic analysis (Mäkitalo et al., 2009) of the dialogical meaning making process of the groups suggests that the students interpret the test questions and problems also from other contexts than those intended by the test-constructors. For instance, words like “pattern” and “diagram” occurring in the test questions can activate a mathematical context that is blurring the intended scientific problem of the task. Illustrations such as graphs and diagrams further mediate such meanings and meaning potentials (Rommetveit, 1974).

Conclusion

In sum, the framing backstories seem to distract the students and to bring about, not situatedness in real-life, but instead uncertainty. In their collaborative meaning making of the items the students appearing in this study enter different contexts in which their discussions take place. In line with Linell’s suggestions (1998) different contexts – scientific, school, mathematical, and everyday life - are at play simultaneously, and these are seen to infer new meanings and framings for what can be considered possible answers.

The agenda for a written test must presuppose that the intended meanings are transferred to the test-takers with no meaning added or removed. However, the students that are working together with the PISA units here, when given the space to discuss the items in order to conclude on a common answer to each question, do not automatically enter the intended contexts. Instead different contexts are activated as the students act as learners of many disciplines at the same time. This raises questions about inferring so called “everyday context” in standardized knowledge testing, and it emphasizes that students are not isolated islands correctly tuned into predetermined wavelengths of particular meanings, but are part of a greater whole in which they act.

References


