

TEACHING AND ASSESSING WITH NEW METHODOLOGICAL TOOLS (MERLO): A NEW PEDAGOGY?

Ferdinando Arzarello¹, Ron S. Kenett², Ornella Robutti³, Uri Shafir⁴,
Theodosia Prodromou⁵, Paola Carante⁶

¹ F. Arzarello, Dept of Math., Univ. of Turin, ferdinando.arzarello@unito.it

² R.S. Kenett, Univ. of Turin and KPA Ltd., ron@kpa-group.com

³ O. Robutti, Dept of Math., Univ. of Turin, ornella.robutti@unito.it

⁴ U. Shafir, KMDI, iSchool, Univ. of Toronto, ushafir@oise.utoronto.ca

⁵ T. Prodromou, School of Education, University of New England, theodosia.prodromou@une.edu.au

⁶ P. Carante, Dept of Math., Univ. of Turin, paola.carante@unito.it

Abstract: The core element of this paper is an innovative didactical and methodological tool, called MERLO (Meaning Equivalence Reusable Learning Objects) used in formative assessment activities. After a general presentation of the MERLO approach, we will focus on the design process of MERLO items for the teaching and learning of mathematics, especially in secondary schools. The paper presents several methodological choices in the setting up and deployment of the MERLO pedagogy. The data, analysis and results come from the research experience gained at the University of Turin in the context of a master's degree of second level for mathematics teacher educators. This is an ongoing research with a current focus on teachers' professional development, with future perspectives on MERLO implementation in the classroom.

Keywords: MERLO, design, teachers' professional development, mathematics teaching and learning, formative assessment.

1. Introduction

It is well known that mathematical objects are sophisticated cultural products that are accessible only by means of representations (Duval, 2006), that is through suitable semiotic representations: as Duval points out "there is no noesis without semiosis" (Duval, 1995, p. 5, 22: Noesis is the intentional act of intellect, and can be defined as the action and the effect of understanding). This is the main reason why semiotic systems are central to mathematical activities and understanding, as pointed out by many scholars, such as Duval himself, Johnson-Laird (1983), Peirce (1931-1958, 1977), Sfard (2000) and Leung, Graf and Lopez-Real (2006). A first consequence of such a situation is what Duval calls the "cognitive paradox" of access to mathematical objects: "How can [students] distinguish the represented object from the semiotic representation used if they cannot get access to the mathematical object apart from the semiotic representations?" (Duval, 2006, italics in the original). A second consequence is that for grasping the meaning of mathematical objects one must cope with the multiple semiotic representations of the same mathematical object in more than one semiotic register, and with their mutual relationships. These capabilities are fundamental for understanding mathematics and consequently crucial for its effective learning (Duval, 2006). The ability to shift from one representation of an object to another representation of the same object is a competence that students should acquire in order to access the underlying meaning (Duval, 2006). The ability to shift between representations is evaluated in international (PISA, TIMSS) and national assessment tests (INVALSI, in Italy).

Students will be able understand statistical and mathematical concepts when teachers take into consideration the relevance of trans-coding capabilities required by students in trans-coding from one semiotic register to the other (for example from a distribution of data to its equation of the probability distribution), or within the same semiotic register (for example from a histogram of data to another equivalent to it - e.g., a pie chart of data). Nevertheless, such capabilities generally are only an implicit object of teaching: no explicit space is given to them in the teaching design, where generally the attention is more on the content, and no attention is given to whether or not they happen in the classroom when students solve problems. Some difficulties may arise when a student does not grasp the same mathematical object represented by different signs (e.g. the equation $y = x^2$ and the graph of a parabola in the Cartesian plane). These difficulties can be classified as part of the so called “duplication obstacle” (Duval, 1983). The duplication obstacle is a real source of difficulty in the learning of mathematics, as reported in literature (Fishbein, 1987), sometimes ignored or underestimated by teachers and currently difficult to be solved in the practice of the didactics of mathematics. A proposal for didactical interventions that avoid or overcome these obstacles is presented in Arzarello, Kenett, Robutti, Shafir (in press). The extent of conceptual understanding (also called “deep understanding”), representing the ability of students to overcome the duplication obstacle, can be assessed with MERLO (Meaning Equivalence Reusable Learning Objects), a methodology that was developed, tested, and evolved since the 1990s by Uri Shafir and Masha Etkind at the Ontario Institute for Studies in Education (OISE) of the University of Toronto, and Ryerson University in Toronto, Canada (Shafir and Etkind, 2010). Arzarello, Kenett, Robutti, Shafir (in press) propose a new pedagogical design that has a twofold purpose: 1) to provide teachers with methods for teaching students of various ages (from grade 6 to 10) mathematics and statistics and assessing students’ deep understanding of quantitative concepts; and 2) to provide students with tools to determine their level of comprehension of statistical and mathematical concepts in formative assessments.

2. The MERLO approach

2.1 Theoretical framework

MERLO (Arzarello, Kenett, Robutti, Shafir, in press; Etkind, Kenett, Shafir, 2010) is a database that allows the sorting and mapping of important concepts through exemplary target statements of particular conceptual situations, and relevant statements of shared meaning. Specifically, each element in the MERLO database is a structured item, anchored to a target statement that describes a conceptual situation and encodes different features of an important concept; each element also includes other statements that may or may not share equivalence-of-meaning with the target. In a mathematical context, for example, an element of MERLO database could be about “parabola”: then this element could include a target statement with the definition of parabola, and other statements in different kinds of representations (symbolic notation, Cartesian graph, table) that share or not share the meaning with the definition of parabola. The figure below is a template for constructing an item anchored to a single target statement.

TARGET STATEMENT						
Surface Similarity [SS]						
Yes			No			
Q1	SS	Yes	SS	No	Q2	Meaning Equivalence [ME]
	ME	Yes	ME	Yes		
Q3	SS	Yes	SS	No	Q4	
	ME	No	ME	No		No

Figure 1: template for constructing a MERLO item

Statements in the four quadrants of the template - Q1; Q2; Q3; and Q4 - are thematically sorted by their relations to the target statement that anchors the particular item. They are classified by two sorting criteria: Surface Similarity and Meaning Equivalence with respect to the target. At an intuitive level we can say that if the statements contain text in natural language, then by Surface Similarity we

mean same/similar words appearing in the same/similar order as in the target statement; while the term Meaning Equivalence designates a commonality of meaning across several representations. Hence, for example the statements “two plus one” and “two plus three” share Surface Similarity but not Meaning Equivalence, while the statements “two pair” and “2+2” share Meaning Equivalence but not Surface Similarity. If we want to go deeper in the interpretation of the two criteria and in particular of the Meaning Equivalence that is the fundamental basis of MERLO, then we can expand the discussion and refer to other famous scholars. Following Frege and many other scholars (for a review see: Steinbring, 2005), who underline a triadic approach to the mathematical meaning, we distinguish between the sign (Zeichen), the sense (Sinn) and the denotation (Bedeutung) of a mathematical inscription. For example $y = 1$ and $y = \sin^2(x) + \cos^2(x)$ are different for the signs in which they are written and for their senses (a function expressed directly or as a sum of trigonometric functions) but they have the same denotation (that is the constant real function equal to 1). Two statements can be equivalent in the mathematical-logical sense because they denote the same mathematical object, even if their senses can be different; moreover they can be formulated in the same or different registers. For example $y=x^2$ and its graph in a Cartesian plane have the same denotation (and sense) even if they are formulated in different registers (Algebraic Vs Cartesian). This is the essential basis of our definition of Meaning Equivalence (essentially equivalence with respect to the denotation, independently from the sense, the sign and the register). However when signs are involved, namely a semiotic interpretation process is involved, things become more complex since the sense attached to a sign is the product of an interpreting activity by a subject (a student). This activity may require more or less steps of reasoning: for example the equivalence between $y = 1$ and $y = \sin^2(x) + \cos^2(x)$ can be more direct than the equivalence of a more complex trigonometric equality. Because of this we introduce below (in section 4.2) the notion of Boundary of Meaning, which is a finer hierarchy between statements, which are equivalent but more or less “distant” according to the number of steps in the reasoning that are necessary to realize the equivalence.

2.2 Some elements of MERLO pedagogy

A typical MERLO item contains 5 unmarked statements: an unmarked target statement plus four additional (unmarked) statements from quadrants Q2; Q3; and Q4. Creators’ experience shows that inclusion of statements from quadrant Q1 makes the item too easy (Etkind M., Kenett R.S., Shafirir U., 2010); as a consequence Q1 statements are avoided in MERLO items. The figure below shows an example of MERLO item, related to Italian national educational programme m@t.abel: the sheet is about data and forecasts and in particular it deals with frequency distribution of discrete values.

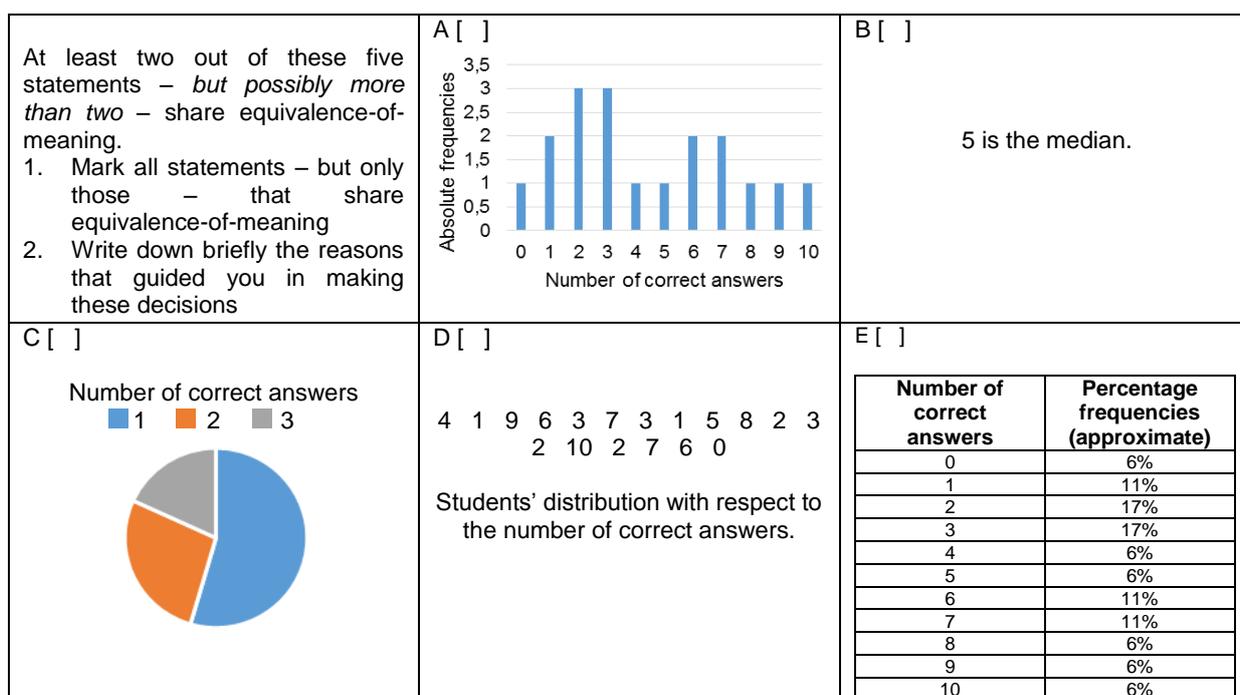


Figure 2: example of MERLO item

As we can see from the example, the learner is required to recognize the statements in multiple representations that share equivalence-of-meaning and to write the concept that he/she had in mind when making the decisions. In this way MERLO items combine multiple-choice (recognition) and short answer (production). It gives a feedback on students with two main scores: recognition and production. The first score comes from the recognition of statements with shared meaning among the given 5 statements, while the second score comes from the writing production of reasons for the decisions. This feedback is useful to the teacher, who can use information about the level of understanding of their students on a particular conceptual knowledge. Specific comprehension deficits can be traced as depressed recognition scores on quadrants Q2 and Q3, due to the mismatch between the valence of Surface Similarity and Meaning Equivalence. The production score of MERLO test items is based on the clarity of learner's description of the conceptual situation anchoring the item, and the explicit inclusion in that description of lexical labels of relevant and important concepts and relations.

Classroom implementation of MERLO pedagogy includes weekly interactive MERLO quizzes, as well as inclusion of MERLO items as part of mid-term tests and final exams. The MERLO interactive quiz is a weekly in-class procedure that provides learners with opportunities to discuss, in small groups, a PowerPoint display of a MERLO item; then send their individual responses to the instructor's computer via mobile text messaging, or by using a clicker (CRS – Classroom Response System). A typical lesson based on the use of MERLO items engages students in: a) small group discussions, b) individual responses, c) class discussion. During the final phase c) the whole class discusses about the given answers and about the reasons for choosing/unchoosing of each of the 5 representations in the MERLO according to the criteria of shared equivalence-of-meaning. This pedagogical approach is very different from the usual classroom scenario where students are given an exercise and are asked to solve it individually. The MERLO approach is structured to provide an opportunity for students to discuss and exchange ideas while thinking about a specific MERLO item; to share and contrast points-of-view; to prompt and refresh each other's memory regarding important details of the conceptual situation; and to 'compare notes' about possible responses.

Regular use of MERLO items in different instructional situations (like for example in architecture programs at the Department of Architectural Science at Ryerson University in Toronto) has been documented to enhance learning outcomes (Shafirir and Etkind, 2006; and Etkind and Shafirir, 2013). For that reason, we recognize the importance of MERLO items as a teaching tool, in order to face a pervasive didactical phenomenon - the *duplication obstacle* - and as a tool to support students in working with multiple semiotic representations and their mutual relationships of the same mathematical object. In this perspective, it is becoming known to teachers through an education programme for in-service mathematics teachers of secondary schools at the University of Turin (master's degree of second level for teacher educators at the Department of Mathematics). In the next section we will describe this experience.

3. Research context and experience description

The research experience we describe here is set in the context of a master's programme of second level for mathematics teacher educators, held at the Department of Mathematics in the University of Turin. It is aimed at in-service Italian secondary-school mathematics teachers; it has 30 enrolled members, who will become educators for other teachers at the end of the courses. The training programme lasts two years and takes place with face-to-face meetings (1 day a week) and distance work. It covers both pedagogical aspects and mathematical content related to the four basic standards that are part of the curriculum in Italy and in many other countries around the world (numbers, geometry, relations and functions, data and forecasts). The aim of this paper is to present experience gained in teachers' professional development with MERLO: we describe this experience and present some important results we reached. We think that this experience can be very useful both at the level of research on MERLO and for its implementation in mathematics teaching, learning and assessing.

During the first year (2013/2014) of programme, teachers were involved in the following training phases:

Phase 1. Translation of MERLO items, produced in other countries (e.g., Russia) and solution of them;
Phase 2. Construction of new MERLO items in geometry, suitable for the Italian school context and curriculum;

Phase 3. Solution of MERLO items produced during phase 2, and analysis of data.

Some findings are available in Arzarello et al. (in press), where the first year of experimentation is presented and analysed.

The research experience inside the programme's context is going on during the second year (2014/2015) with the whole group as learners and practitioners and with a small group of seven voluntary teachers as teacher-researchers. The teacher-researcher is an important figure, because it represents a joining element between University and secondary school contexts. From September 2014 to March 2015 these seven teacher-researchers worked closely with researchers, focusing in particular on aspects related to the design of MERLO items. We would like to highlight the fact that researchers and teachers have different backgrounds, knowledge and experiences, but they share working together. As result, we arrived to some shared methodological choices about the task design of MERLO items, we will present in the next section. Collection of data was carried out through notes, minutes and some videos of the face-to-face meetings; e-mail conversations and shared materials for distance work.

4. Experience analysis: how to design a MERLO item

The aim of this section is to present the general process of design that can be followed and which leads to a complete MERLO item. In particular we show some shared methodological choices, emerged from the research experience (especially during the second year of the master's programme) and from the analysis of the collected data.

The starting point for the construction of a new MERLO item is to identify a relevant concept on which to base the whole activity. Some questions might arise, like the following: when can a concept be considered 'relevant' enough to be explored with a MERLO? And what is a good way to organize the mathematical knowledge domain into relevant concepts? In order to solve these issues, the community of researchers suggested to the community of teachers to design MERLO items starting from m@t.abel activities and INVALSI tests (well known by Italian teachers), where the mathematical knowledge is already organized into four main conceptual nodes (numbers, geometry, relations and functions, data and forecasts) and into more specific topics and relations between concepts. Therefore the relevant concepts on which to focus MERLO items are borrowed from m@t.abel and INVALSI. We reiterate that m@t.abel is a national educational programme, launched in 2006 in Italy to significantly impact quantitative literacy competencies in the Italian schools; while INVALSI is the national assessment test in Italy for detecting and measuring students' learning at different levels of instruction.

4.1 Target Statement

The immediately following step is to create a Target Statement (TS) connected to the relevant concept borrowed from m@t.abel or INVALSI. Experience with teacher-researchers showed that this step is far from trivial, because it requires a transformation from a traditional kind of question (posed as open question or problem) into a less traditional form. The common practice to present a problem in form of question is reflected on the first MERLO items produced by teachers at the beginning of the experimentation, where the TS was written as open question. After some meetings with researchers and MERLO revisions by the whole group, teachers acquired the practice of elaborating the TS as a statement, graph, or table.

4.2 Q2 statements

An essential point in the design of MERLO items is the creation of a Boundary of Meaning associated to the chosen Target Statement: the boundary includes statements that share Meaning Equivalence with the Target Statement; hence Q2 statements are inside this boundary. We think that the introduction and the reflection on the Boundary of Meaning helped us very much in order to clarify what we mean saying that two or more statements share Meaning Equivalence. Indeed Meaning Equivalence is a very delicate matter, especially in a mathematical context (in section 2.1 we gave our interpretation of it).

At the beginning of the experience, the majority of the Q2 statements produced by teachers were the result of what that Duval calls conversions:

Conversions are transformations of representation that consist of changing a register without changing the objects being denoted: for example, passing from the algebraic notation for an equation to its graphic representation, passing from the natural language statement of a relationship to its notation using letters, etc. (Duval R., 2006, p.112)

However restricting MERLO items to Duval's conversions limits the use and the applicability of MERLO for the teaching and learning. Indeed from teachers' didactical point of view, other aspects are also important to foster students' learning, like giving examples and showing the application of mathematical concepts in different contexts. This is what emerged during a discussion at a meeting (December 3th, 2014):

Teacher L. *"If I want to make sure that students understood the various writings, I have to give them an example, otherwise ..."*

Teacher A. (who teaches both mathematics and physics): *"In physics examples are necessary. I prepared a sheet on the derivatives, where in addition to the definition, I also used examples of application in physics."*

These are meaningful words representing some teachers' needs that led to the decision to include also examples and application of concepts to problems inside the boundary of shared meaning. Of course the different types of Q2 statements, although within the same boundary, are more or less "distant" to the TS. As mentioned above in section 2.1, the "distance" depends on the number of steps in the reasoning that are necessary to realize the equivalence between Q2 and TS. We think that the positioning of the Q2 statements into three different levels is suitable to classify them according to the distance from the TS (the choice of three levels seems reasonable, as they may be sufficient without being excessive). If the number of steps in the reasoning is too large, we can consider the statement very distant and so out of the boundary. The statements classification into levels may be useful for diversified assessment: higher or lower score depending on the level of difficulty in recognizing statements of shared meaning.

4.3 Other statements

We have just described the design of TS and Q2 statements; now we are going to examine the other kinds of statements. First of all we restate that in general the use of Q1 statements is avoided: creators' experience shows that inclusion of statements from quadrant Q1 makes the item too easy (Etkind M., Kenett R.S., Shafrir U., 2010).

In relation to design of Q3 and Q4 statements, researchers and teachers considered two different possibilities: they always have to be mathematically true by themselves (not compared in pairs) or they can be not true, including a widespread mistake or misconception, common among students. This is not a simple choice because it strongly influences the kind of test and the information that you get on students' knowledge and skills. Discussions on this choice are numerous, recurring also over time. Eventually the group agreed on the methodological choice that all statements must be mathematically true, because reasoning about true meanings is more formative from a didactical point of view: indeed MERLO is designed to promote conceptual thinking. The justification for the choice transpires from the following words, written during an e-mail conversation (December 10th, 2014):

Teacher S.: *"We need to decide exactly what we want to highlight in students' preparation. [...] If we put Q3 and Q4 as not true statements, then in my view we change the objective, while it should be on meanings. [...] Students' misconceptions, perhaps, still emerge from MERLO with well designed Q3 and Q4."*

Another important aspect emerged from the experience on the design process of a MERLO item is the following: if Q3 and Q4 statements are created only in relation to the TS on the bases of Meaning Equivalence and Surface Similarity, then they could share a meaning among them. We want to avoid Q3 and Q4 statements about the same concept and link together, otherwise there might be two

Boundaries of Meaning and two possibilities to complete the task: one marking TS and Q2 and the other marking Q3 and Q4. In the perspective of using MERLO for assessment we prefer a single answer. The formulation of Q3 and Q4 statements requires to refer to the TS but also requires a comparison between them. Teachers have to put attention on this aspect during the design process and they have to highlight it to students, saying them that each MERLO item has a single answer.

4.4 Task for students

The original MERLO task for students, formulated in English and readable in the example of MERLO item in figure 2, required a rethinking for adapting it to a mathematical context inside an Italian culture. After detailed discussions among researchers and teachers, the task for students assumes the following Italian formulation:

1. *Segnare le affermazioni che condividono lo stesso significato (due o più);*
 2. *Indicare le ragioni che guidano nella scelta.*
- [1. Mark the statements that share the same meaning (two or more);
2. Write the reasons that guided you in the choice.]

4.5 Example

Here is an example designed by a teacher of the group, in which all the shared methodological choices discussed above are made evident. In particular we can see the reformulation of the task for students, the formulation of the TS as statement, the mathematical exactness of all statements, the Boundary of Meaning associated to the TS, that includes also a Q2 statement representing an application of the concept in a different context (in physics). Therefore this example shows the versatility of the tool, also for interdisciplinary links.

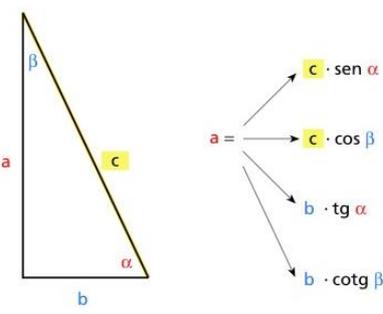
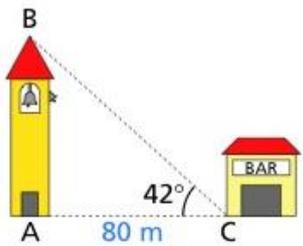
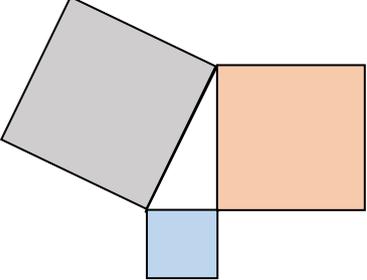
<p>1. Mark the statements that share the same meaning (two or more);</p> <p>2. Write the reasons that guided you in the choice.</p>	<p>TS A []</p> <p>In a right-angled triangle the measure of a cathetus is equal to:</p> <ul style="list-style-type: none"> • the measure of the hypotenuse multiplied by the sine of the opposite angle with respect to the cathetus or by the cosine of the adjacent angle with respect to the cathetus • the measure of the other cathetus multiplied by the tangent of the opposite angle with respect to the first cathetus or by the cotangent of the adjacent angle with respect to the first cathetus 	<p>Q2 B []</p> 
<p>Q2 C []</p>  <p>The height of the bell tower is 72 m</p>	<p>Q2 D []</p> <p>The intensity and direction of the resultant of the two forces of intensity $F_1 = 8 \text{ N}$ and $F_2 = 10 \text{ N}$ applied at point A and forming an angle of 30° is $F_R = 17,39 \text{ N}$</p>	<p>Q4 E []</p> 

Figure 3: example of MERLO item designed by a teacher

5. Conclusion and future perspectives

In this paper, we looked at the use and design of MERLO items, in particular the experience gained in MERLO design from the programme participants. Another interesting and ongoing research project is related to MERLO implementation in classroom. MERLO is an effective tool for formative assessment, because it allows frequent feedback about students' understanding. Therefore it may be useful to integrate it in teaching practice, combining MERLO with the use of other modern didactical tools, like education software. We can mention for example TinkerPlots, a software for dynamic data exploration developed for use by middle school through university students (www.tinkerplots.com), or GeoGebra, a dynamic mathematics software for teaching and learning, which links geometric and algebraic aspects (www.geogebra.org). Both of them expose students to a different learning experience through an alternative approach to statistics or geometry, based on multiple representations and visualizations. MERLO is in line with this underlying idea and so it presents itself as a suitable tool for formative assessment, also inside less traditional learning paths.

References

- Arzarello, F., Kenett, R.S., Robutti, O., Shafrir, U. (in press). The application of concept science to the training of teachers of quantitative literacy and statistical concepts.
- Duval, R. (1983). L'obstacle du dedoublement des objets mathematiques, *Educational Studies in Mathematics*, vol. 14 pp. 385-414.
- Duval, R. (1995). *Semiosis et pensée humaine. Registres sémiotiques et apprentissages intellectuels*. Bern: Peter Lang.
- Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics*, 61, 103-131.
- Etkind, M., Kenett, R.S., Shafrir, U. (2010). The evidence based management of learning: diagnosis and development of conceptual thinking with meaning equivalence reusable learning objects (MERLO). In: *8th International conference on teaching statistics (ICOTS)*, Ljubljana, Slovenia.
- Etkind, M. and Shafrir, U. (2013). Teaching and Learning in the Digital Age with Pedagogy for Conceptual Thinking and Peer Cooperation. International Association of Technology, Education and Development. *Proceedings of INTED-2013 Conference*, Valencia, March 2013, pp. 5342-5352.
- Fischbein, E. (1987). *Intuition in Science and Mathematics, an Educational Approach*, D. Reidel Publishing Company: Dordrecht.
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, interference and consciousness*. Cambridge, UK: Cambridge University Press.
- Leung, F.K.S., Graf, K., and Lopez-Real, F. (2006). Mathematics Education in Different Cultural Traditions- A Comparative Study of East Asia and the West: *The 13th ICMI Study*. Springer
- Peirce, C.S. (1931-1958). *Collected Papers (Volumes I-XIII)*. Eds. Harthstorne, P.W.C., and Weiss, P. Cambridge (Ma): Harvard University Press.
- Peirce, C.S. (1977). *Semiotics and Significs*. Ed. Hardwick C., Bloomington (IN): Indiana University Press.
- Sfard, A. (2000). Symbolizing mathematical reality into being: How mathematical discourse and mathematical objects create each other. In P. Cobb, K. E. Yackel, & K. McClain (Eds), *Symbolizing and communicating: perspectives on Mathematical Discourse, Tools, and Instructional Design* (pp. 37-98). Mahwah, NJ: Erlbaum.
- Shafrir, U. and Etkind, M. (2006). eLearning for depth in the semantic web. *British Journal for Educational Technology*, 37(3), 425-444.
- Shafrir, U. and Etkind, M. (2010). *Concept Science: Content and Structure of Labeled Patterns in Human Experience*. Version 31.0.
- Steinbring, H. (2005). *The Construction of New Mathematical Knowledge in Classroom Interaction – An Epistemological Perspective*, Mathematics Education Library, vol. 38, Berlin: Springer.