ENHANCING LEARNING OUTCOMES WITH 'BIG DATA' FROM PEDAGOGY FOR CONCEPTUAL THINKING WITH MEANING EQUIVALENCE REUSABLE LEARNING OBJECTS (MERLO) AND INTERACTIVE CONCEPT DISCOVERY (INCOD)

Masha Etkind[1], Theodosia Prodromou[2], and Uri Shafrir[3]
[1]Ryerson University, Toronto, Canada, [2]University of New England, Armidale, Australia
[3]University of Toronto, Canada
metkind@ryerson.ca

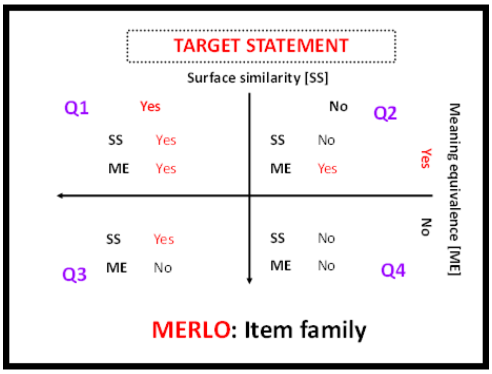
Learning outcomes of traditional pedagogy focus on memory of facts, problem-solving procedures, and multiple-choice or true/false questions. Pedagogy for conceptual thinking focuses on higher-order thinking skill, exploration of equivalence of meaning among ideas and relationships between statements that denote commonality of meaning across representations. MERLO assessments capture these important aspects of conceptual thinking. 'Big data' from the results of weekly formative MERLO quizzes reveal individual students' conceptual strengths and weaknesses, showing details of the evolution of deep understanding of each concept in the course. This allows the instructor to suggest corrective measures with Interactive Concept Discovery (InCoD) in the course' digital Knowledge Repository (KR), conducted and discussed by individual students with their peers, and enhance learning outcomes in large undergraduate courses in statistics and mathematics.

WHAT IS MERLO?

Higher-order conceptual thinking skills are now recognized as a cornerstone of effective learning, ways of thinking that explore patterns of meaning in the context of a conceptual framework (Bransford, Brown, & Cocking. 2004). Meaning equivalence is a construct that denotes common meaning across representations, a polymorphous, one-to-many transformation of meaning. Meaning Equivalence Reusable Learning Objects (MERLO) is a pedagogical tool for both teaching and assessment, by asking students to identify important concepts through exemplary target statements of particular conceptual situations, and relevant statements that may – or may not – share that same meaning. MERLO assessment items for different concepts, combined into a large database for a course of study, can provide valuable information about student learning patterns (Etkind, Kenett, & Shafrir, 2010; 2016). Each node in such a MERLO database is an item family, anchored by a Target Statement (TS) that describes a conceptual situation and encodes different features of an important concept; and also includes other statements that may or may not share equivalence-of-meaning with TS.

Collectively, these item families encode the full conceptual content of a course (or a particular content area within a discipline). Statements are sorted into the four quadrants of the template by their relation to the Target Statement that anchors the particular node (item family) using two sorting criteria: Surface Similarity (SS) to the target; and Meaning Equivalence (ME) with the target (Figure 1). 'Surface similarity' refers to similar modes of representation, e.g., use of natural language, or use of a graph, and similar elements, such as similar or the same words in different orders as the target statement; and by 'meaning equivalence' we mean that, in a community that shares a sublanguage (Kittredge, 1983) with a controlled vocabulary (e.g., the community of statistics educators), a majority would agree that the meaning of the statement being sorted is equivalent to the meaning of the target statement. Each MERLO assessment item contains 5 unmarked, and randomly arranged, statements: an unmarked Target Statement plus four additional (unmarked) statements from quadrants Q2; Q3; and Q4. Our experience has shown that inclusion of statements from quadrant Q1 makes a MERLO assessment item too easy, because it gives away the shared meaning due to the valence-match between surface similarity and meaning equivalence, a strong indicator of shared meaning between Q1 and the target statement. Therefore, Q1 statements are excluded from MERLO assessment items (see Figure 2).

Figure 1: Template for constructing an item-family in MERLO



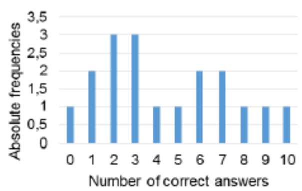
Task instructions for MERLO assessment are: "At least two out of these five statements – *but possibly more than two* – share equivalence-of-meaning. (1) Mark all statements – but only those – that share equivalence-of-meaning. (2) Write down briefly the reasons for making these decisions." The first task is a *recognition task*: the target statement is not marked, so features of the concept to be compared are not made explicit. In order to perform this task, a learner needs to decode and recognize the meaning of each of the 5 statements in the set (A; B; C; D; E). This decoding process is carried out, typically, by analyzing concepts that define the 'meaning' of each statement. Successful analysis of all the statements requires deep understanding of the conceptual content of the specific domain. The MERLO item format requires both rule inference and rule application in a similar way to the solution of analogical reasoning items. Once the learner marks those statements that – in his/her opinion – share equivalence-of-meaning, he/she formulates and briefly describes in writing the concept (i.e., idea/criteria) he/she had in mind when making these decisions, which is considered a *production task*. Figure 2 is an example of a MERLO assessment item (Statistics), with 5 representations, in text, numbers, colors, tables, and diagrams.

Figure 2: Example of a MERLO assessment item (Statistics)

At least two out of these five statements – *but possibly more than two* – share equivalence-of-meaning.

1. Mark all statements – but only those – that share equivalence-of-meaning
2. Write down briefly the reasons that guided you in making these decisions

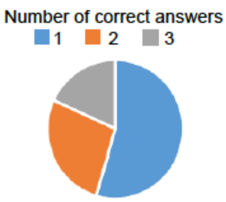
A[ ]



B[ ]

5 is the median.

C[ ]



D[ ]

4 1 9 6 3 7 3 1 5 8 2 3 2 10 2 7 6 0

Students' distribution with respect to the number of correct answers.

E[ ]

Number of correct answers	Percentage frequencies (approximate)
0	6%
1	11%
2	17%
3	17%
4	6%
5	6%
6	11%
7	11%
8	6%
9	6%
10	6%

BOUNDARY OF MEANING (BOM): METRICS FOR DIAGNOSTICS OF MISCONCEPTIONS

There are two scores for each MERLO item, the *recognition score* and *production score*. These two scores provide clear and reliable evidence for diagnosing misconceptions, and clues for remediation. The scores help teachers understand how individual students are drawing their boundary of meaning in contrast to the boundary of meaning as set by the larger statistics community. Given a community of specialists that share a sublanguage, and a Target Statement that encodes a particular conceptual situation; then *'Boundary of Meaning' (BoM)* is defined as *the boundary between two mutually exclusive semantic spaces* in the sublanguage:

- a semantic space that contains *only representations that share equivalence-of-meaning* with the Target Statement.
- a semantic space that contains *only representations that do not share equivalence-of-meaning* with the Target Statement.

Figure 3: Boundary of Meaning (BoM) of a particular TS include Q2 statements; and exclude Q3 statements and Q4 statements

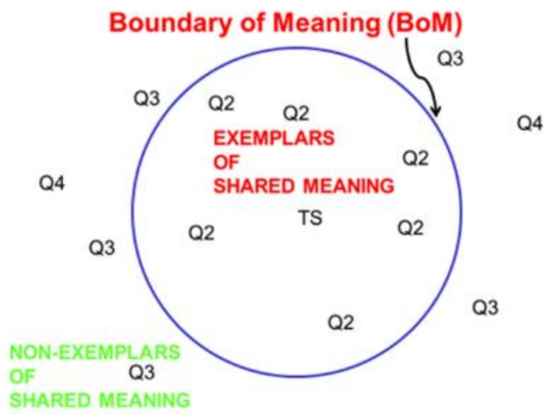


Figure 3 illustrates the demarcation of the BoM for a particular Target Statement in a knowledge domain. Specific conceptual comprehension deficits are traced if the learner's response to a MERLO assessment item with 5 unmarked statements, *includes choosing Q3 statements*; as well as if the learner *does not choose Q2 statements*. Therefore, 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, as shown in Figure 1 above.

Figure 4 shows mean class scores for TS; Q2; Q3; and Q4; of twelve MERLO assessment items in final exam of 11-grade physics/magnetism course at Russian Academy of Sciences, Ioffe Physico Technical Institute, Lycee Physico Technical High School, St. Petersburg (2003). Proportion correct scores for each student were calculated for Target Statement TS, and quadrants Q2, Q3, and Q4; for example:

$$\text{Q2 score} = \frac{\text{Number of } \textit{correctly chosen} \text{ Q2 statements in the twelve MERLO assessment items}}{\text{Total number of Q2 statements in the twelve MERLO assessment items}}$$

$$\text{Q3 score} = \frac{\text{Number of } \textit{correctly unchosen} \text{ Q3 statements in the twelve MERLO assessment items}}{\text{Total number of Q3 statements in the twelve MERLO assessment items}}$$

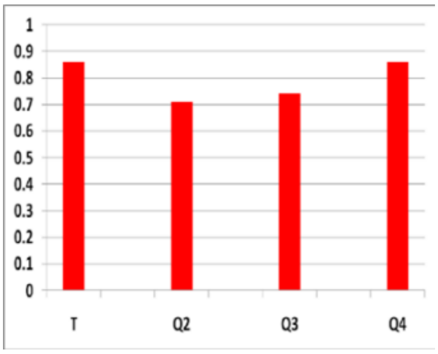
The U-shaped form of these results reveals reduced scores in Q2 and Q3 representations. However, the interpretations of these two reduced scores are very different:

- *Reduced score on Q2* indicates that the learner *fails to choose* (include in the BoM of the concept) certain statements that do share equivalence-of-meaning (but do not share surface similarity) with the target; such depressed Q2 score signals *an over-restrictive (too exclusive)* understanding of the meaning underlying the concept.

- *Reduced score on Q3* indicates that the learner did choose (*fails to exclude* from the BoM of the concept) certain statements that do not share equivalence of-meaning (but that do share surface similarity) with the target; this depressed Q3 score signals *an under-restrictive (too inclusive)* understanding of the meaning of the concept.

  Therefore, the U-shaped form of the results in Figure 4 reveals certain *deficiencies of both over-restrictive and under-restrictive understanding of the meaning of the concept.*

Figure 4: MERLO final exam mean recognition scores for T and quadrants Q2, Q3, and Q4; (physics/magnetism), Lycee Physico Technical High School, St. Petersburg (2003).



WEEKLY QUIZZES WITH MERLO FORMATIVE ASSESSMENTS

Pedagogy for conceptual thinking was designed to motivate and engage learners, and encourage cooperation. It includes weekly classroom formative assessment with a MERLO item for one important concept, and provides learners with opportunities to cooperate through discussions in small groups and to make their own decisions, sending their individual responses to the instructor's computer via mobile communication devices. These activities are followed by class discussion. MERLO formative assessment takes about 20 minutes, and includes the following 4 steps:

- *Peer cooperation and small group discussion* (approx. 3-5 minutes), includes: PowerPoint projection of the MERLO item; student discussion of the item in small groups (turn back/sideways to discuss with those seated next to them).
- *Individual response* (3 min); each student sends his/her *individual recognition response* through a personal communication device (smartphone/tablet/clicker), marking at least 2 out of 5 statements (A; B; C; D; E) in the multi-semiotic MERLO item that, in his/her opinion, share equivalence-of-meaning; then sends his/her *individual production response*, briefly describing the conceptual reasons he/she had in mind when making the above recognition decisions.
- *Feedback and class discussion of students' production responses* (3-5 min), including PowerPoint projection of the MERLO item with several production responses sent by individual students, and the instructor's description of the MERLO conceptual context; class discussion and comparison of the various individual production responses.
- *Feedback and class discussion of students' recognition responses* (3-5 min), including PowerPoint projection of the MERLO item, showing the correct recognition responses; class discussion and comparison of the various individual recognition responses.

Data collected in weekly formative MERLO quizzes reveal the depth of conceptual understanding - and the specific misunderstandings - of each student with respect to the course content.

INTERACTIVE CONCEPT DISCOVERY (InCoD)

In addition to weekly MERLO formative quizzes, pedagogy for conceptual thinking encourages learners to conduct weekly Interactive Concept Discovery (InCoD) searches in the course

Knowledge Repository (KR) that include collection of digital documents related to course content (e-books; journals; e-databases; reports; etc.). InCoD is a novel semantic search learning tool, developed, tested and implemented, based on Concept Parsing Algorithms (CPA; Shafrir & Etkind, 2018). It is an intuitive, interactive procedure that allows a learner to search large digital databases, and to discover the building blocks of a concept within a particular context of the course knowledge domain, namely, co-occurring subordinate concepts and relations. InCoD constructs concept maps that clearly identify not only the conceptual content of important concepts in course material, but also its internal conceptual structure - hierarchical and lateral relations among concepts and their building blocks. The learner begins the process of Interactive Concept Discovery by conducting semantic search (concordance) of Key Word In Context (KWIC); then reads and saves found relevant documents; annotates and evaluates the degree of relevance of a particular document to the specific conceptual content under consideration; and constructs graphical representations of links between concepts. InCoD data reveal the learner's consistency of 'drilling-down' for discovering deeper building blocks of the particular concept, as well as the temporal evolution of outcomes of the learning sequence. This 'big data' digital record is an authentic, evidence-based demonstration of acquisition of mastery of knowledge that can be used as a springboard for follow-up classroom and chat-room discussions. It provides a credible record to the individual's learning process and learning outcomes (Shafrir, Etkind, & Treviranus, 2006).

Interactive Concept Discovery (InCoD) makes available to the learner all the different locations of sentences in the digital documents in the course Knowledge Repository (KR), likely written by different authors with different points-of-views and different examples that contain a particular searched concept. Clicking on a found sentence provides access to the document where the sentence appears, so the learner can examine the context in which the concept is discussed. Following the initial concordance/search, the learner may notice that in several locations/sentences, another concept consistently appears in close proximity to the searched concept. By clicking on the second concept, the learner activates a subsequent *search of co-occurrence of both the initial concept and the second concept*, aimed at discovering consistent co-occurrence with the initial concept. A *Learner Individual Index* records data from each of the learner's activities, including:

- Alphabetic indexing by name of concept
- Document/page
- Ranking (on scale of 1 to 5) of degree of relevance to course content
- Annotations, including the learner's brief summary of the specific conceptual context, and other comments, tags, and links

Clicking on a particular entry in an Individual Index provides the date of creation of the entry by the learner, plus complete details of content. A Learner Individual Index is also available to the instructor, which tracks the learner's progress in mastery and conceptual understanding of the documents in the Knowledge Repository, and the specific learning outcomes accumulated throughout the course.

Concepedia (Conceptual Encyclopedia) is a weekly aggregation of all Individual Indexes of all learners in the class, in the context of the course knowledge domain, and is accessible to the instructor and to all students in the course. It also includes learners' commentaries on other learners' annotations. *Concepedia enhances individuals' reputations as cooperators who contribute to the public good, and reflects the cumulative process-learning-curve of the class.*

'BIG DATA' EVIDENCE-BASED MERLO AND INCOD LEARNING ANALYTICS

Detailed data of learning processes and outcomes are collected, analyzed and available to the learners and to the teachers through learning analytics (Shafrir & Etkind, 2006; Shafrir & Kenett, 2016). Each individual student's 'big data' profile shows the student's Individual Index of InCoD, as well as scores of MERLO weekly formative assessments for specific concepts, and mid-term tests and final exams in the courses. These scores identify specific deficits in conceptual understanding of course content, expressed as lower individual Q2 and Q3 scores, and document corrective interventions with individual learners.

Class data profiles show mean MERLO scores in weekly formative assessments, mid-term tests and final exams, as well as *Concepedias* for different courses. These data indicate class-specific deficits, expressed as lower mean Q2 and Q3 class scores in conceptual understanding of a particular course conceptual content, and may prompt the instructor to revisit this content in future lectures or other class activities. These inter-related learning-process data are collected continuously, not just in a particular class, but across all learning and teaching activities throughout the semester and the academic year. Eventually, and subject to strict privacy procedures that protect individual student identity and privacy, this learning analytics is available to the administration of the academic institutions.

CONCLUSION

A recent OECD review provides strong evidence for the important role of formative assessment in enhancing students' learning outcomes (Nusche, 2013). The present paper discusses enhancing learning outcomes and deep understanding of concepts through interactive learning and methods for assessing such understanding. In the digital age, when 'big data' is abundant and technology is within reach of everyone, the focus on depth of understanding is gaining importance and urgency. This paper demonstrates advantages of learning in the context of modern information technologies. The wide range of experience in MERLO and InCoD tools demonstrates their universality and reflects their large potential in future research. Implementation, testing, and validation, since 2002, of pedagogy for conceptual thinking, lend support to the following conclusions:

- Weekly MERLO formative assessments enhance peer cooperation and conceptual thinking.
- Weekly Interactive Concept Discovery (InCoD) provides individual learners with different points of views of conceptual situations relevant to the course content, written by different authors.
- Pedagogy for conceptual thinking motivates and engages students. This is particularly evident - and important - in *large undergraduate classes in statistics and mathematics*.
- Conceptual thinking is learneable.
- Pedagogy for conceptual thinking, when implemented as a regular part of the instructional methodology, replicates the above pattern of results and enhanced learning outcomes.

REFERENCES

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2004). *How People Learn: Brain, Mind, Experience, and School* (expanded edition). Washington, DC: National Academy Press.

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). Invited paper. *Proceedings of the 8th International Conference on Teaching Statistics* (ICOTS8). Ljubljana, Slovenia.

Etkind, M., Kenett, R. S., & Shafrir, U. (2016). Learning in the Digital Age with Meaning Equivalence Reusable Learning Objects (MERLO). *Handbook of Research on Applied Learning Theory and Design in Modern Education* (Vol. 1, pp. 310-333). IGI Global.

Kittredge, *R.I.*, (1983). Semantic Processing of Texts in Restricted Sublanguages. *Computational Linguistics.* Cercone N.J. (Ed.), 45-58. Oxford: Pergamon.

Nusche, R. (2013). Student assessment: Putting the learner at the centre. In *Reviews of Evaluation and Assessment in Education and Assessment. Synergies for Better Learning: An International Perspective on Evaluation* (pp. 139-270). Paris: OECD Publishing.

Shafrir, U., & Etkind, M. (2006). eLearning for Depth in the Semantic Web. *British Journal for Educational Technology, 37(3),* 425-444.

Shafrir, U., & Etkind, M. (2018). *Concept Parsing Algorithms (CPA) for Textual Analysis and Discovery: Emerging Research and Opportunities.* ISBN13: 9781522521761. IGI Global.

Shafrir, U., Etkind, M., & Treviranus, J. (2006). eLearning Tools for ePortfolios. *Handbook of Research on ePortfolios*, pp. 206-216. Hershey: IGI Global. doi: 10.4018/978-1-59140-890-1.ch020

Shafrir, U., & Kenett, R. S. (2016). Concept Science Evidence-Based MERLO Learning Analytics. *Handbook of Research on Applied Learning Theory and Design in Modern Education*. *Volume 1, Chapter 16,* 334-357. ISBN 9781466696341. IGI Global.