



**Giornale Italiano della Ricerca Educativa**

**Italian Journal of Educational Research**

**RIVISTA SEMESTRALE**

anno XII – numero speciale MAGGIO 2019

**Guest editors Valentina Grion, Anna Serbati, Ettore Felisatti, Lan Li**



**Direttore | Editor in chief**

PIETRO LUCISANO | Sapienza Università di Roma

**Condirettori | Co-editors**

LOREDANA PERLA | Università degli Studi di Bari "Aldo Moro"

ACHILLE M. NOTTI | Università degli Studi di Salerno

**Comitato Scientifico | Editorial Board**

JEAN-MARIE DE KETELE | Université Catholique de Lovanio

VITALY VALDIMIROVIC RUBTZOV | City University of Moscow

MARIA JOSE MARTINEZ SEGURA | University of Murcia

ETTORE FELISATTI | Università degli Studi di Padova

GIOVANNI MORETTI | Università degli Studi di Roma Tre

ALESSANDRA LA MARCA | Università degli Studi di Palermo

ROBERTO TRINCHERO | Università degli Studi di Torino

LORETTA FABBRI | Università degli Studi di Siena

IRA VANNINI | Università di Bologna

ANTONIO MARZANO | Università degli Studi di Salerno

MARIA LUISA IAVARONE | Università degli Studi di Napoli "Parthenope"

GIOVANNI BONAIUTI | Università degli Studi di Cagliari

MARIA LUCIA GIOVANNINI | Università di Bologna

PATRIZIA MAGNOLER | Università degli Studi di Macerata

ELISABETTA NIGRIS | Università degli Studi Milano-Bicocca

**Comitato editoriale | Editorial management**

CRISTIANA DE SANTIS | Sapienza Università di Roma

DANIA MALERBA | Sapienza Università di Roma

ROSA VEGLIANTE | Università degli Studi di Salerno

**Note per gli Autori | Notes to the Authors**

I contributi, in formato MS Word, devono essere inviati all'indirizzo email del Comitato Editoriale: [rivista@sird.it](mailto:rivista@sird.it)

Ulteriori informazioni per l'invio dei contributi sono reperibili nel sito [www.sird.it](http://www.sird.it)

---

Submissions have to be sent, as Ms Word files, to the email address of the Editorial Management: [rivista@sird.it](mailto:rivista@sird.it)

Further information about submission can be found at [www.sird.it](http://www.sird.it)

**Consultazione numeri rivista**

<http://ojs.pensamultimedia.it/index.php/sird>

**Codice ISSN 2038-9744 (testo on line)**

**Registrazione Tribunale di Bologna n. 8088 del 22 giugno 2010**

**Finito di stampare: MAGGIO 2019**

**Abbonamenti • Subscription**

Italia euro 25,00 • Estero euro 50,00

Le richieste d'abbonamento e ogni altra corrispondenza relativa agli abbonamenti vanno indirizzate a: [abbonamenti@edipressrl.it](mailto:abbonamenti@edipressrl.it)

**Editing e stampa**

Pensa MultiMedia Editore s.r.l. - Via A. Maria Caprioli, 8 - 73100 Lecce - tel. 0832.230435  
[www.pensamultimedia.it](http://www.pensamultimedia.it) - [info@pensamultimedia.it](mailto:info@pensamultimedia.it)

**Progetto grafico copertina**

Valentina Sansò

## Obiettivi e finalità | Aims and scopes

Il **Giornale Italiano della Ricerca Educativa**, organo ufficiale della **Società Italiana di Ricerca Didattica (SIRD)**, è dedicato alle metodologie della ricerca educativa e alla ricerca valutativa in educazione.

Le aree di ricerca riguardano: lo sviluppo dei curricoli, la formazione degli insegnanti, l'istruzione scolastica, universitaria e professionale, l'organizzazione e progettazione didattica, le tecnologie educative e l'e-learning, le didattiche disciplinari, la didattica per l'educazione inclusiva, le metodologie per la formazione continua, la docimologia, la valutazione e la certificazione delle competenze, la valutazione dei processi formativi, la valutazione e qualità dei sistemi formativi. La rivista è rivolta a ricercatori, educatori, formatori e insegnanti; pubblica lavori di ricerca empirica originali, casi studio ed esperienze, studi critici e sistematici, insieme ad editoriali e brevi report relativi ai recenti sviluppi nei settori. L'obiettivo è diffondere la cultura scientifica e metodologica, incoraggiare il dibattito e stimolare nuova ricerca.

---

The **Italian Journal of Educational Research**, promoted by the **Italian Society of Educational Research**, is devoted to Methodologies of Educational Research and Evaluation Research in Education.

Research fields refer to: curriculum development, teacher training, school education, higher education and vocational education and training, instructional management and design, educational technology and e-learning, subject teaching, inclusive education, lifelong learning methodologies, competences evaluation and certification, docimology, students assessment, school evaluation, teacher appraisal, system evaluation and quality.

The journal serves the interest of researchers, educators, trainers and teachers, and publishes original empirical research works, case studies, systematic and critical reviews, along with editorials and brief reports, covering recent developments in the field. The journal aims are to share the scientific and methodological culture, to encourage debate and to stimulate new research.

## Comitato di referaggio | Referees Committee

Il Comitato di Revisori include studiosi di riconosciuta competenza italiani e stranieri. Responsabili della procedura di referaggio sono il direttore e il condirettore della rivista.

---

The Referees Committee includes well-respected Italian and foreign researchers. The referral process is under the responsibility of the Journal's Editor in Chief and Co-Editors.

## Procedura di referaggio | Referral process

Il Direttore e Condirettore ricevono gli articoli e li forniscono in forma anonima a due revisori, tramite l'uso di un'area riservata nel sito della SIRD ([www.sird.it](http://www.sird.it)), i quali compilano la scheda di valutazione direttamente via web entro i termini stabiliti. Sono accettati solo gli articoli per i quali entrambi i revisori esprimono un parere positivo. I giudizi dei revisori sono comunicati agli Autori, assieme a indicazioni per l'eventuale revisione, con richiesta di apportare i cambiamenti indicati. Gli articoli non modificati secondo le indicazioni dei revisori non sono pubblicati.

**Per consultare il codice etico consultare il link:**

<http://ojs.pensamultimedia.it/index.php/sird/about/editorialPolicies#custom-0>

---

Editor in chief and co-editor collect the papers and make them available anonymously to two referees, using a reserved area on the SIRD website ([www.sird.it](http://www.sird.it)), who are able to fulfill the evaluation grid on the web before the deadline. Only articles for which both referees express a positive judgment are accepted. The referees evaluations are communicated to the authors, including guidelines for eventual changes with request to adjust their submissions according to the referees suggestions. Articles not modified in accordance with the referees guidelines are not accepted.

# INDICE

## 9 EDITORIALE

**VALENTINA GRION, ANNA SERBATI, ETTORE FELISATTI & LAN LI**

La valutazione all'Università: riflessioni dal passato e prospettive per il futuro | *Assessment and evaluation at University: reflections from the past and future perspectives*

## Studi

## 15 ETTORE FELISATTI

La valutazione all'Università: riflessioni dal passato e prospettive per il futuro | *Assessment and evaluation at University: reflections from the past and future perspectives*

## 29 DEBORA AQUARIO

Quale valutazione per l'apprendimento? Verso la promozione di una assessment identity | *Which Assessment for learning? Towards the promotion of an assessment identity*

## 45 KAY SAMBELL, SALLY BROWN & PHIL RACE

Peer feedback and technology-enhanced assessment as critical issues to foster student learning | *Peer feedback e technology-enhanced assessment come fattori chiave per favorire l'apprendimento degli studenti*

## 63 LAN LI

ENGAGE students in formative peer Assessment to support learning | *ENGAGE: coinvolgere gli studenti nella valutazione formativa tra pari per supportare l'apprendimento*

## 71 DAVID NICOL

Reconceptualising feedback as an internal not external process | *Riconcettualizzare il feedback come processo interno e non esterno*

## 85 MASSIMO MARCUCCIO & LILIANA SILVA

Peer feedback as assessment practice in doctoral programs: a systematic review of empirical research | *La valutazione tra pari come pratica valutativa nei dottorati di ricerca: una rassegna sistematica delle ricerche empiriche*

## 101 LUCIANO GALLIANI

Tecnologie e valutazione: bio-bibliografia di un intreccio | *Technology and assessment: bio-bibliography of an integrated story*

## Ricerche

- 115 ANNA SERBATI, VALENTINA GRION, MARINELLA FANTI**  
Caratteristiche del *peer feedback* e giudizio valutativo in un corso universitario blended | *Peer feedback features and evaluative judgment in a blended university course*
- 139 DONATELLA CESARENI & NADIA SANSONE**  
Il *peer-feedback* collaborativo per il miglioramento continuo dei prodotti | *The collaborative peer-feedback for the continuous improvement of products*
- 157 ELISA TRUFFELLI & ALESSANDRA ROSA**  
Peer-feedback individuale e di gruppo: uno studio empirico sull'utilità percepita in un corso universitario sulla valutazione | *Individual versus group peer feedback: an empirical study on perceived usefulness in a Higher Education course on Evaluation*
- 177 LAURA FOSCHI & GRAZIANO CECCHINATO**  
Validity and reliability of peer grading in in-service teacher training | *Validità e affidabilità del peer-grading nella formazione di insegnanti in servizio*
- 195 EMILIA RESTIGLIAN & VALENTINA GRION**  
Valutazione e feedback fra pari nella scuola: uno studio di caso nell'ambito del progetto GRiFoVA | *Assessment and peer feedback in school contexts: a case-study carried out by GRiFoVA group*
- 223 MASSIMO MARGOTTINI & FRANCESCA ROSSI**  
Strumenti per l'autovalutazione di competenze strategiche per lo studio e il lavoro | *Self-assessment tools for strategic learning and vocational skills*
- 241 ANTONIO MARZANO & SERGIO MIRANDA**  
Dynamic concept maps to support e-learning and evaluation processes | *Mappe concettuali dinamiche per supportare i processi di e-learning e di valutazione*
- 257 CORRADO PETRUCCO**  
Student Response Systems as a successful tool for formative assessment: students' perceptions in a university pilot study | *Student Response Systems come strumento per la valutazione formativa: la percezione degli studenti in un pilot study*

267 **MICHELE BIASUTTI**

Forum e wiki a confronto come strumenti per l'apprendimento collaborativo online | *Forums and wikis as tools for online collaborative learning: a comparative study*

## Esperienze

291 **ALESSIA BEVILACQUA**

Un'esperienza di valutazione formante in ambito universitario: l'attenzione ai Moodle Logs per capire i tempi e misurare le frequenze d'uso di una piattaforma e-learning da parte degli studenti | *Assessment as Learning in higher education. Promoting active learning in large classes through educational technologies*

# Dynamic concept maps to support e-learning and evaluation processes

## Mappe concettuali dinamiche per supportare i processi di e-learning e di valutazione

Antonio Marzano • Department of Human, Philosophical and Educational Sciences • University of Salerno (Italy)  
Sergio Miranda • Department of Human, Philosophical and Educational Sciences • University of Salerno (Italy)

When technologies enter the learning processes, it is always advisable to analyse the changes they trigger carefully. This is because their effects are difficult to predict and generalize. There are several factors related to the times, ways, status, culture, even digital, of the subjects involved and therefore any statement should derive from careful observations and cautious analysis. This article moves along the line of research traced by a laboratory of the University of Salerno, the laboratory RIMEDI@, using for a further academic year, a technology within a learning environment, the e-Lena platform. In fact, this use derives from a first phase of experimentation from which the results have emerged identified possible future developments. After that, a series of corrective measures have been implemented to proceed now with this second experiment, which is either to confirm or refute the results already obtained and to verify the effectiveness of the changes made. The access to the environment is provided in addition to the classroom teaching, so it assumes a form of support and not a substitute form of in presence teaching. We tried to verify if our enriched online learning environment could stimulate remediation processes between digital and analogic artefacts in students and represent a mean to support automatic evaluation processes. In other words, we would answer the question related about if it favours the development of meaningful learning. The use of this technology was therefore observed and the data collected analysed to be able to synthesize conclusions and outline new objectives in order to move the next steps in research. The results confirm once again that the environment built in e-Lena offers an effective support to students. It takes a significant role both in activating remediation experiences and in automating evaluation processes, so it could encourage the development and the consolidation of learning.

**Keywords:** Assessment, e-Learning, Dynamic Concept Maps, Automatic test generation

Quando le tecnologie “entrano” nei processi di apprendimento, è opportuno analizzare con attenzione i cambiamenti che attivano. Questo perché gli effetti che producono sono difficili da prevedere e da generalizzare. Sono diversi i fattori che interagiscono (tempi, modi, cultura anche digitale dei soggetti coinvolti) e ogni affermazione dovrebbe derivare da osservazioni accurate e analisi prudenti. Questo articolo si muove lungo una direttrice di ricerca del laboratorio RIMEDI@ dell’Università di Salerno nel quale, nell’ultimo triennio, è stato sperimentato ed implementato un ambiente virtuale di apprendimento all’interno della piattaforma e-Lena. L’accesso all’ambiente viene fornito in aggiunta all’insegnamento in aula, quindi assume una forma di supporto e non una forma sostitutiva di insegnamento in presenza. Obiettivo dell’indagine è stato quello di verificare se, e in qual misura, l’ambiente di apprendimento online potesse stimolare negli studenti processi di rimediazione tra artefatti digitali e analogici e rappresentare un mezzo per supportare processi automatici di valutazione. I risultati confermano che l’ambiente costruito in e-Lena offre un supporto efficace agli studenti. Esso assume un ruolo significativo sia nell’attivazione delle esperienze di rimediazione che nell’automazione dei processi di valutazione, in modo da incoraggiare lo sviluppo e il consolidamento degli apprendimenti.

**Parole chiave:** Valutazione, e-Learning, Mappe concettuali dinamiche, Generazione automatica di test


241

ricerche

The authors shared the concept of the work and elaborated the text according to the following breakdown. Antonio Marzano, scientific coordinator of the research, provided the proposition and the general lines of the work, being for this indicated as first name; he also took care of the sections Introduction, 1. Theoretical context and objectives, 4. Data analysis. Sergio Miranda took care of the sections 2. Methodology, 3. The TEST-Maker tool. Both authors elaborated the section 5. Conclusions.

# Dynamic concept maps to support e-learning and evaluation processes

## Introduction



The survey described in this paper has been carried out in the academic year (AY) 2017-18 and represents a further development of a broader experimental plan launched in the AY 2014-15. The objective of the first research was to verify if and how it was possible to experiment with learning evaluation methods supported by automatic test generation processes able to provide diagnostic information on the progress of learning (formative evaluation), to favour self-assessment, offer feedback for improvement in progress of the training paths designed by the teacher. We were starting from a prototype developed by Miranda et al. (2013) which, from a knowledge domain and a semantic modelling based on ontologies (Sutt & Motta, 2004; Naeve et al., 2006), was able to produce test batteries related to all domain topics and to the logical relations between them (Alsubait, Parsia, Sattler, 2012). The tool (*TEST-Maker*) had been experimented in the teaching of “IT and telematic skills” (18 hours, 3 CFU) of the degree course in “Education of playrooms, socio-educational laboratories and recreational services for the childhood” held in the second semester of the AY 2014-15 at the University of Salerno. The battery of tests generated by *TEST-Maker*, delivered to 89 students, had been validated in terms of its overall reliability and the results emerged empirically confirmed the correctness of the conceptual and algorithmic structure at the base of the automatic generation tool (Marzano, Miranda, & Orciuoli, 2017). In the subsequent academic years (2015-16, 2016-17), *TEST-Maker* had been further improved and implemented.

The empirical survey carried out in the AY 2016-17 was designed as a further “segment” of the multi-year experimental plan which we referred to at the beginning of the paragraph and with the aim of responding to a specific question: can a specifically designed environment stimulate reciprocal remediation processes in students (Bolter & Grusin, 2002) between digital and analogical artefacts and encourage the development of meaningful learning, one in interaction and hybridization with the other one? The survey involved 201 students enrolled in the bachelor of “Science in Primary Education” at the University of Salerno attending the “School and Educational Design” course. The experimental data were collected by elaborating and using four tests for the verification of learning. The network interaction space



for experimentation was the *e-LENA* platform, the Learning Management System of the RIMEDI @ research laboratory. The analogical artefacts we consider are mainly books, papers and any kind of material that is not available on digital devices.

The results that emerged (Marzano, 2017) were positive and confirmed our hypotheses: the virtual environment was an effective support for the students, a device that encouraged the processes of remediation between digital and analogical artefacts, favoured the development and the consolidation of learning.

The references presented up to now constitute the starting point of the research carried out in the academic year 2017-18 and of which we are going to present, after a theoretical and methodological study, some specific processes and aspects.

## 1. Theoretical context and objectives

This research is part of the use of technologies in teaching, with the aim of verifying the impact in terms of learning. Specifically, we analysed the synergistic use of dynamic concept maps (hereinafter DCM) in an e-learning platform to verify the activation of remediation processes by students attending the AY 2017-18. Methods and techniques of educational research (Master's degree in "Professional educators and experts in continuing education") of the University of Salerno. In particular, we wanted to verify if the DCM could stimulate processes of reflection, the necessary in-depth analysis (also on the book) and thus succeed in bringing a benefit compared to the conventional study alone. The attempt was to understand its formative potentials by proposing a possible interpretation of this process by using the historical-cultural approach of Vygotsky (1974), enriched and "contaminated" by the perspective of reciprocal remediation of Bolter and Grusin (2002) inspired to the theories of McLuhan (1967). "A medium is what remedies" (Bolter & Grusin, 2002); the remediation is the way to "refashion" a media. For instance, photography remediated painting, internet sites remediated newspaper pages. In our study, we consider the remediation processes related to the didactic resources for our courses and the new digital way to keep their knowledge available.

In order to trigger remediation, we assumed that it is essential to have a well-designed learning environment. The interaction space for experimentation was that of the *e-LENA* platform created by customizing the *Moodle* solution and integrating it by means of other components specifically developed in order to respond to the needs associated with the research initiative.



The Vygotsky historical-cultural approach is fundamental to our understanding of the relationships between artefacts and minds. For Vygotsky (1974), the learning and psychic development of the person are determined by biological, instrumental, historical and socio-cultural factors. The cultural context, through its signs and tools, acts as a mediator of the cognitive process of the person. The use of an instrument contributes to the formation of concepts that transcend the instrument itself and therefore allows the emergence of meanings and the construction of new knowledge to interact and solve problems. The artefact can act as a mediator between the individual and the world, triggering transformations in the cognitive characteristics of the individuals themselves (Bonaiuti, 2005). Using an artefact, as a result of intentional and non-involuntary processes (Rossi & Toppano, 2009), transforms the same knowledge for which it was designed (Norman, 1991) and in its design we must therefore consider the task for which it will be used.



The DCMs that we tested in this work are cognitive artefacts developed to favour the processes of conceptual systematization through a graphical representation that, according to Novak and Gowin (1989), can favour the emergence of the meanings inherent in the materials to learn and stimulate the reflection on the nature of knowledge and on the relationships between them. Visualizing knowledge through concepts and logical links allows better understanding it, to select the main information compared to the secondary ones and to trigger processes of retrospective analysis and re-elaboration (Rivoltella, 2010). In the field of teaching, when tools for the representation of knowledge are used, two recurring limits concern their static nature and their being limited to broad but generic or specific and restricted domains. These objective difficulties prompted us to design the DCMs as a mediation artefact that allow us navigating the concepts by opening the nodes and discovering from time to time the nodes of existing children and relationships, as well as accessing to the contents associated with them. The DCMs, in our research, have been designed to verify if, and to what extent, they could stimulate reciprocal reconstruction processes between digital and analogical artefacts in both directions (Bolter & Grusin, 2002) and, consequently, favour the hybridization between different mediums (Latour, 1991) for the development of meaningful learning.

These are the basic theoretical references for the research that has been proposed to investigate the role of learning environments in the network, meaning them as devices built with regulatory values. We have made a change, compared to the experimentation carried out in the previous year: the tests were processed by using *TEST-Maker*. The aim of this work is therefore to verify the reliability of the tests automatically generated by using *TEST-Maker*.

## 2. Methodology

The students attending the course of “Methodologies and techniques of the educational research” of the Master’s Degree in “Professional educators and experts in continuing education” (30 hours, 6 CFU) were involved.

The educational objectives of teaching are to promote the ability to develop and implement research activities, also encouraging the development of practical-applicative skills compared to some techniques and tools that can be used in the different phases of the experimental plan. The teaching program includes, during the assessment of the learning, a mid-term (optional) and a written preliminary and mandatory test to support the interview. The written exam consists of 30 multiple-choice questions divided into two sections (the first 15 items related to the first training segment of the course and the subsequent items, from number 16 to 30, relating to the second segment) and is considered to have been exceeded with at least 8 replies exact for each of the two sections of the test. With the passing of the intermediate exam (whose verification refers to the first training segment) the students, in the subsequent written test (the obligatory one to hold the interview), must answer only the last 15 items of the test.

The educational activities started on October 10th (ending November 30th 2017). In the first two lessons (10th and 12th October), the research project was presented and voluntary participation in the experimentation was proposed.

The second training segment of the course consisted of the study of three chapters taken from one of the two recommended texts for individual study. Participants would be informed during the previous lesson (on Tuesday) the chapter to be studied to take the written test to be held on the following Thursday. The test, consisting of 15 multiple-choice questions, was passed by answering exactly at least 8 questions. The passing of the three tests was equivalent to passing the second section of the written exam. Passing also the optional intermediate test, the students could directly access the oral interview. The experimentation consisted of dividing the participating students into two groups by sampling technique: the students of the first group, the control group (CG), used the book for the preparation for the test; the students of the second group, the experimental group (EG), used the *e-LENA* platform. 59 students participated in the experimentation. From the list, sorted alphabetically, by systematic sampling, the two groups were defined (Table 1).



	CG	EG
<b>Males</b>	3.3(1)	6.9 (2)
<b>Females</b>	96.7 (29)	93.1 (27)
<b>Total</b>	100	100
<b>(N)</b>	30	29

**Tab. 1: The composition of the two groups**

The trial began on November 7th and the first test was delivered on November 9th. The subsequent tests were carried out on a weekly basis (November 16th and 23rd). The study materials (Table 2) were taken from the book *Methodology of research in education and training* (2002).



Date	Materials
November 9th	The experimental design
November 16th	The evaluation survey
November 23th	Evaluation tools

**Tab. 2: The calendar of the experimentation and the study materials**

The concept maps were developed according to the Quillian (1968) approach. For each topic, the concepts of a higher order have been identified, from which other less general concepts are expanded hierarchically, which are articulated in further categorization (Marzano & Miranda, 2018). The concepts were placed in a semantic relationship between them (Lalumera, 2009) through the use of predicates (Galliani & Notti, 2014). Three different dynamic maps were constructed, containing respectively 54, 59 and 56 concepts. The maps created were added to the *e-LENA* platform as educational resources. When accessing these resources, the user-student displays only the root node, can click and display all the relationships that start from this node and the related “children” nodes connected to it. On the displayed nodes, you can repeat the opening actions of the relationships and the child nodes or access the contents associated with each of them. The nodes and relationships, in fact, are displayed one level at a time leaving the user free in navigation and allowing him to customize his experience of discovery (Fig. 1).

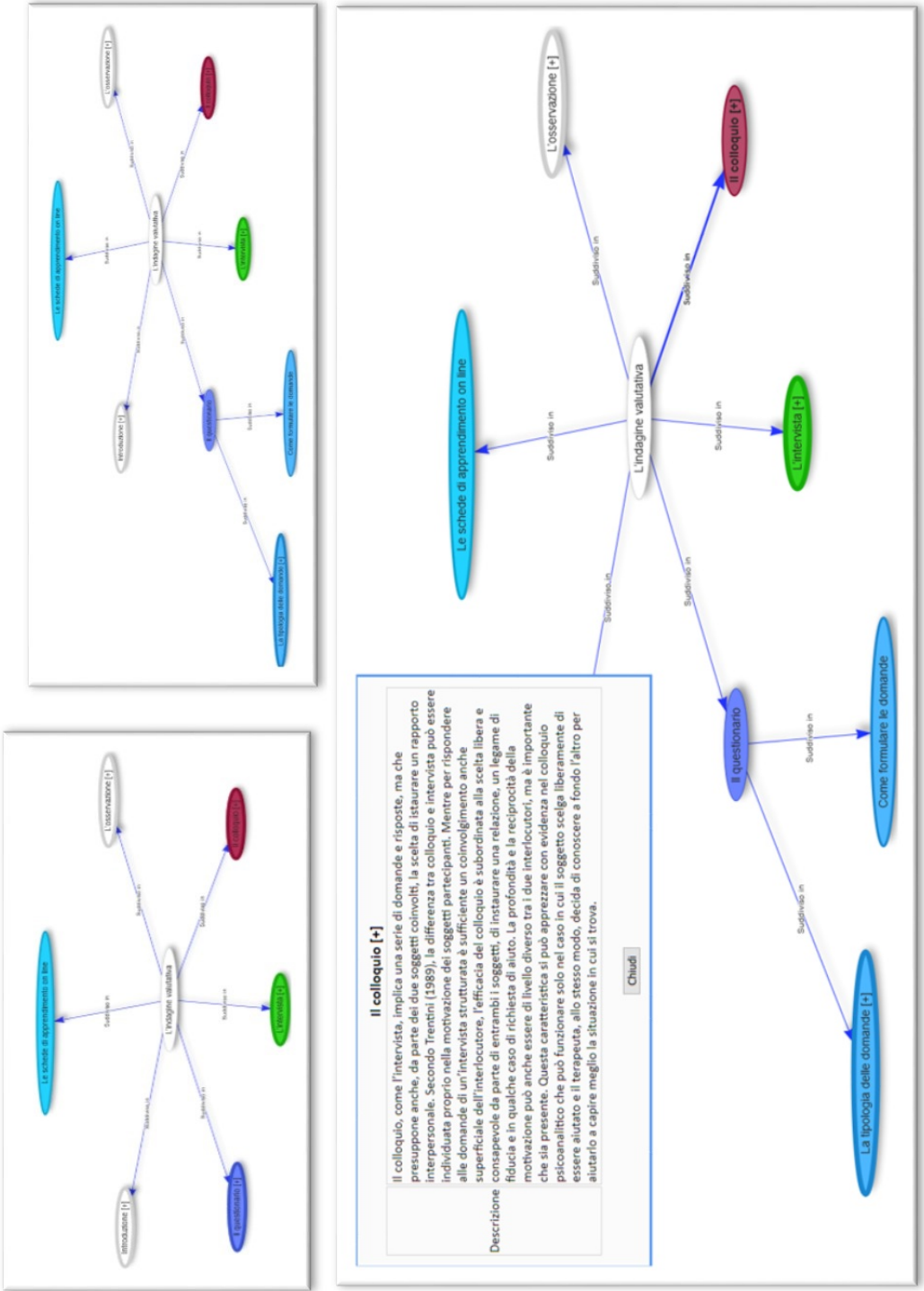


Fig. 1: The dynamic concept map

### 3. The *TEST-Makertool*

The automatic generation of tests is based on the design of a concept map as a useful tool for the purposes of a common and shared organization of a specific knowledge domain. The schemes used refer to the semantic web paradigm whose architecture aims to enrich data and information with metadata and semantic maps to facilitate understanding and processing. Considering the definition of Gruber (1993), according to which a semantic map is the specification of a conceptualization (through an appropriate definition/description of the concepts and relationships that may exist for an object or a set of objects), we can affirm that they facilitate the interpretability of content and constitute a basis of “learning objects” organized through the terminological consent for their indexing and designed with the aim of being reused and shared. The use of maps makes it possible to identify explicitly determined relationships by defining knowledge structures and inference rules.

Having said that, we considered a concept map as a shared knowledge structure negotiated and formalized by the scientific community through semantic description languages (Studer et al., 1998). The ontological structure reflects the shared conceptual core and aims to integrate the formal dimension of a knowledge domain with the relational one of the daily activities of the different communities of practices and learning. In essence, they are configured as support environments for the representation of knowledge: they organize the key concepts of a specific knowledge domain, define and interconnect them dynamically between them through semantic relations (Marzano & Notti, 2015). Conceptual networks thus formalized want to qualify as instruments mediators of the complexity of a field of knowledge.

In the case under examination, the concepts used for the definition of the three maps refer to the topics and themes reported in Table 2. Exploiting the procedures of Euzenat and Shvaiko (2013) according to which an ontological representation with any number of relationships can be brought back to one in which the types of relationships present are only relations of order or decomposition, it was chosen to draw the concept map (also known as the subject ontology in Miranda et al., 2016) with only three types of relations (*Subdivided into*, *Prerequisite of*, *Suggested Order*) in order to simplify the representation of the domain of knowledge taken into consideration and therefore to reduce the algorithmic complexity during the generation.

Starting from these ontologies, our tool is able to generate the questions for the verification tests automatically. In order to articulate its composition, some “distractors” are also created using concepts and descriptions external to the treated domain as well as non-existent relationships.



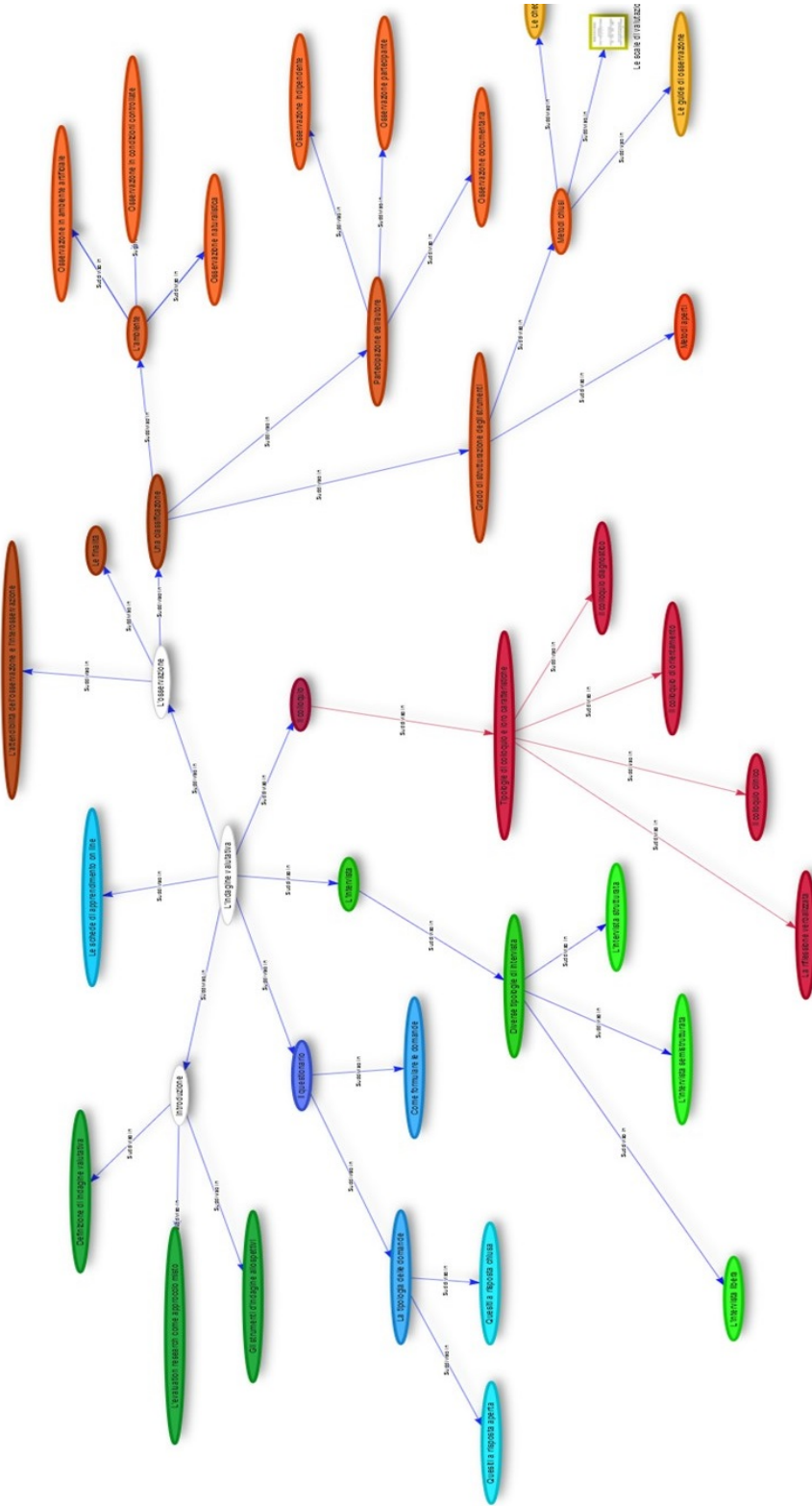


Fig. 2: The concepts and the relations of the map "L'indagine valutativa"



The types of questions generated (in our case we have used only the multiple choice questions) can only be textual or with images in the body of the application itself. The text is created with simple rules of composition using the name of the concept, the description, the outgoing relationships and the names of the concepts connected by these relationships. The images are created either by taking parts of the ontology or by adopting different criteria of representation (ensembles, analogies, puzzles).

The logic of creating the questions is mainly linked to the navigation of the relationships of the knowledge map. The level of complexity of the questions can vary according to the depth of the relationships that goes from the simplest direct link (level 1) to more complex semantic subsumed relations or links outlined by a chain of relations of the same type (level 2, 3, 4...  $n$ , based on the length of the chain).

Once the ontology has been defined, it is possible to “activate” the test generation tool which analyses all the direct and semantically derived relationships and, based on a series of parameters and options, creates the questions to construct the verification tests (Fig. 3).

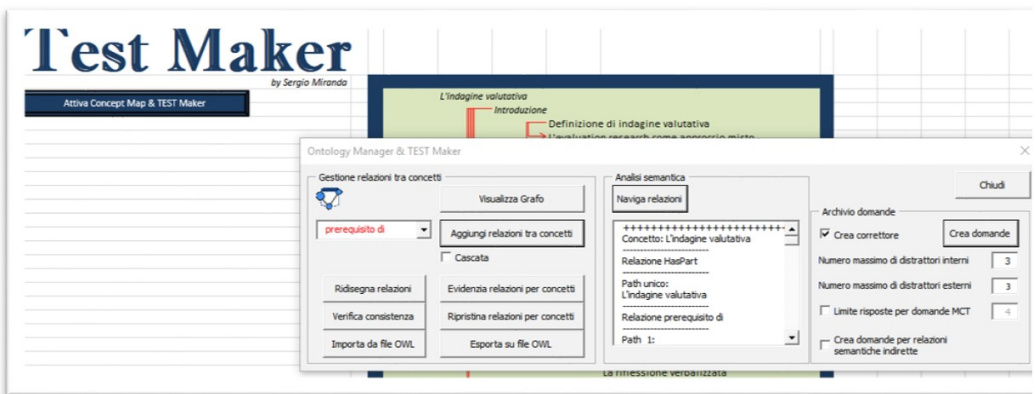


Fig. 3: The Automatic Test Generation Tool

The creation phase is automatic and, based on the choices made, the tool creates a single battery of questions or various random combinations to be used for verification tests to be printed or distributed through, for example, an e-learning platform. An example is shown below (Fig. 4).



TEST generato in automatico sull'ontologia con radice "L'indagine valutativa"

Domande

Introduzione	Il questionario	L'intervista
Il colloquio		Le schede di apprendimento online

1) Quale dei seguenti concetti è rappresentato dalle parti visualizzate in figura?

Genesi  
 Poesia  
 L'indagine valutativa  
 Corriere

Definizione di indagine valutativa → L'evaluation research come approccio misto  
 prerequisite di

2) L'immagine rappresenta la relazione tra i due concetti. Quale delle seguenti interpretazioni è corretta?

Il concetto "Definizione di indagine valutativa" va studiato dopo del concetto "L'evaluation research come approccio misto"  
 Il concetto "Definizione di indagine valutativa" va studiato prima del concetto "L'evaluation research come approccio misto"  
 Il concetto "Definizione di indagine valutativa" va studiato insieme al concetto "L'evaluation research come approccio misto"  
 Lo studio del concetto "Definizione di indagine valutativa" è obbligatorio, mentre quello del concetto "L'evaluation research come approccio misto" è facoltativo

Definizione di indagine valutativa	L'evaluation research come approccio misto	Gli strumenti d'indagine atipici
------------------------------------	--	----------------------------------

3) Quale dei seguenti concetti è rappresentato dalle parti visualizzate in figura?

L'evaluation research come approccio misto  
 Definizione di indagine valutativa  
 Accelerazione  
 Introduzione

Definizione di indagine valutativa → La tipologia delle domande  
 prerequisite di

3) L'immagine rappresenta la relazione tra i due concetti. Quale delle seguenti interpretazioni è corretta?

Per comprendere il concetto "Definizione di indagine valutativa" serve il concetto "La tipologia delle domande"  
 I concetti "La tipologia delle domande" e "Definizione di indagine valutativa" vanno affrontati insieme  
 Per comprendere il concetto "La tipologia delle domande" serve il concetto "Definizione di indagine valutativa"  
 Comprendere il concetto "Definizione di indagine valutativa" è obbligatorio mentre comprendere il concetto "La tipologia delle domande" è facoltativo

Come formulare le domande	La tipologia delle domande
---------------------------	----------------------------

4) Quale dei seguenti concetti è rappresentato dalle parti visualizzate in figura?

Il questionario



Fig. 4: Examples of tests generated automatically

## 4. Data analysis

The first test (15 items generated by *TEST-Maker*) was given on November 9th to the 59 students participating in the trial. In order to be able to grasp the significance of the collected data, we first proceeded to perform a first series of statistical elaborations (Table 3): the calculation of the central tendency indices (*arithmetic mean*, *modal value*, *Mo* and *median*, *Me*), of the standard deviation and of the Cronbach coefficient.

		<b>Mo</b>	<b>Me</b>		
<b>CG</b>	7.21	8	7	1.97	0.79
<b>EG</b>	7.45	7	7	2.11	

**Tab. 3: The statistics of the first test**

Overall, the test was neither too easy nor too difficult. The value of the standard deviation indicates a fair level of homogeneity of the scores and a distribution of responses almost symmetrically concentrated around the value of the mean as regards the trend of the scores in the two groups. The value of Cronbach's  $\alpha$  indicates a fair internal homogeneity of the test. We present the indices emerged from the processing of data related to the administration of the second test (Table 4).



		<b>Mo</b>	<b>Me</b>		
<b>CG</b>	7.86	8	7	1.67	0.74
<b>EG</b>	7.72	8	7	1.75	

**Tab. 4: The statistics of the second test**

Also in this case, the test was neither too easy nor too difficult. The value of the standard deviation indicates a good level of homogeneity of the scores and the value of the Cronbach  $\alpha$  shows a discrete internal homogeneity of the test. The statistics on the data of the third test are in the Table 5.

		<b>Mo</b>	<b>Me</b>		
<b>CG</b>	8.05	9	8	1.44	0.83
<b>EG</b>	8.21	8	8	1.52	

**Tab. 5: The statistics of the third test**

The value of the standard deviation indicates a good level of homogeneity of scores and a distribution of responses almost symmetrically concentrated around the value of the mean. The value of Cronbach's  $\alpha$  indicates a good correlation and internal homogeneity of the items of which the survey instrument is composed. Overall, the data show a substantial homogeneity about the performance of the students of the two samples.

With the analysis of the questions, some characteristics of the ques-

tions generated by *TEST-Maker* were verified. In particular, the focus was on checking the ability of the tests to adequately select the students according to their level of preparation. In this sense, one of the methods that can be used to ascertain the reliability of the questions used to verify the knowledge held by the students, is the *item analysis* or analysis of the answers, which allows obtaining information about the reliability of a test on the whole and on the functioning of the single items (Notti, 2003). This procedure is carried out to ascertain the degree of difficulty of individual questions and their ability to discriminate in relation to the overall outcome. It is therefore possible to check whether, and to what extent, there are questions that are too trivial or too difficult, ambiguous or ill-structured questions, if each item is discriminating, i.e. able to quantitatively differentiate students based on the different knowledge possessed.

The Table 6 shows the values assumed by the indices of difficulty ( $D$ ) and of discriminativity ( $d$ ) related to the entire battery of applications used, highlighting the criticalities in bold.



	<b>D</b>	<b>d</b>
<b>1</b>	0,5	0,6
<b>2</b>	0,6	0,7
<b>3</b>	0,7	0,6
<b>4</b>	<b>0,3</b>	<b>0,3</b>
<b>5</b>	0,6	0,7
<b>6</b>	0,7	0,7
<b>7</b>	0,6	0,4
<b>8</b>	0,6	0,5
<b>9</b>	0,7	0,8
<b>10</b>	0,6	0,5
<b>11</b>	0,6	0,7
<b>12</b>	<b>0,4</b>	<b>0,3</b>
<b>13</b>	0,6	0,7
<b>14</b>	0,5	0,6
<b>15</b>	0,6	0,5
<b>16</b>	0,7	0,6
<b>17</b>	<b>0,3</b>	<b>0,4</b>
<b>18</b>	0,6	0,7
<b>19</b>	0,7	0,6
<b>20</b>	0,6	0,7
<b>21</b>	0,5	0,7
<b>22</b>	0,5	0,6
<b>23</b>	0,6	0,5
<b>24</b>	0,5	0,6
<b>25</b>	0,6	0,7

26	0,7	0,6
27	0,5	0,7
28	<b>0,8</b>	<b>0,4</b>
29	0,6	0,6
30	0,5	0,7
31	0,8	0,7
32	0,7	0,6
33	0,5	0,4
34	0,6	0,5
35	0,7	0,6
36	0,6	0,7
37	0,5	0,7
38	0,5	0,6
38	0,5	0,6
40	<b>0,4</b>	<b>0,5</b>
41	0,7	0,6
42	0,7	0,6
43	0,6	0,5
44	0,6	0,6
45	0,5	0,6

Tab. 6: The values of the indices D and d

There are five questions that present critical values in the indexes (4, 12, 17, 28 and 40). However, the test is on average reliable and of fair quality. Ultimately, considering what emerged from the data processing, the tools built for the verification of the learning can be considered valid and reliable. This data seems to confirm empirically the correctness of the conceptual and algorithmic structure that is at the base of the automatic test generation tool named *TEST-Maker*.

## 5. Conclusions

The research was a continuation of the survey already carried out in 2016. Although starting from the same objectives and using the same methodology, a modification concerning the elaboration of the tests was introduced using a tool called *TEST-Maker*. In this sense, the objective was to verify the reliability of the tests processed through the use of this semi-automatic generation software with the aim of providing an effective, valid and reliable support tool. While considering that any measure is essentially an approximation (De Luca & Lucisano, 2011), the analysis of the results relative to the quality of the single items and of the overall test is decidedly encouraging and can be, in

perspective, the premise for the development of an application to support the processes of self-evaluation also useful for in progress re/planning the training paths.

The results that emerged must be taken into account in order to launch further and more in-depth investigations. In fact, the latest version of *TEST-Maker* can and must be further improved. For this reason, the authors have planned some actions to be launched during the next AY2019-20 in order to make it even more reliable and effective. Moreover, in parallel, the system will be experimentally integrated as a self-assessment tool in *e-LENA*, the e-learning platform of the *RIMEDIA*, the research laboratory in Media Education and Active Didactics of the Department of Human Sciences, Philosophy and Education of the University of Salerno.

## References

- Alsubait T., Parsia B., & Sattler U. (2012). Automatic generation of analogy questions for student assessment: An ontology-based approach. *Research in Learning Technology*, 20 (suppl.), pp. 95-101.
- Bolter J.D., & Grusin R. (2002). *Remediation. Competizione e integrazione tra media vecchi e nuovi*. Milano: Guerini e Associati (ed. orig. 1999).
- Bonaiuti G. (2005). *Strumenti della rete e processo formativo. Uso degli ambienti tecnologici per facilitare la costruzione della conoscenza e le pratiche di apprendimento collaborative*. Firenze: Firenze University Press.
- De Luca A.M., & Lucisano P. (2011). Item analysis between model and reality. *Giornale Italiano della Ricerca Educativa*, IV, 7, pp. 85-96.
- Euzenat J., & Shvaiko P. (2013). *Ontology Matching*. Springer Science & Business Media.
- Galliani L., & Notti A.M. (Eds.), (2014). *Valutazione educativa*. Lecce: Pensa MultiMedia.
- Gruber T. R. (1993). A translation approach to portable ontologies. *Knowledge Acquisition*, 5(2), pp. 199-220.
- Lalumera G. (2009). *Cosa sono i concetti?* Bari: Laterza.
- Latour B. (1991). *Nous n'avons jamais été modernes*. Paris: Éditions La Découverte (trad. it. *Non siamo mai stati moderni*, Elèuthera, Milano, 1995).
- Lucisano P., & Salerno A. (2002). *Metodologia della ricerca in educazione e formazione*. Roma: Carocci.
- Marzano A. (2017). Mappe concettuali dinamiche e processi di rimediazione in ambienti di apprendimento in rete. *Formazione & Insegnamento*, 3, pp. 245-278.
- Marzano A., & Miranda S. (2018). Mappe dinamiche per “navigare la conoscenza. In *Exploring the Micro, Meso and Macro. Navigating between*



*dimensions in the digital learning landscape European Distance and E-Learning Network*, EDEN 2018 Annual Conference, pp. 843-856.

- Marzano A., Miranda, S., & Orciuoli, F. (2017). L'uso di ontologie per la generazione automatica di un test. In P. Magnoler, A. Notti, & L. Perla (Eds.), *La professionalità degli insegnanti. La ricerca e le pratiche* (pp. 677-695). Lecce: Pensa MultiMedia.
- Marzano A., & Notti A.M. (2015). Eduonto: An Ontology For Educational Assessment. *Journal Of E-Learning And Knowledge Society*, 11, pp. 69-82.
- McLuhan M. (1967). *Gli strumenti del comunicare*. Milano: Il Saggiatore. (Edizione originale pubblicata 1964).
- Miranda S., Loia V., Orciuoli F., Mangione G., & Gaeta M. (2013). *Automatic Generation of Assessment Objects and Remedial Works for MOOCs*. Proceedings of 12th International Conference on Information Technology Based Higher Education and Training, ITHET 2013, art. no. 6671018.
- Miranda S., Orciuoli F., & Sampson D.G. (2016). A SKOS-based framework for Subject Ontologies to improve learning experiences. *Computers in Human Behavior*, 61, pp. 609-621.
- Naeve A., Lytras M., Nejdil W., Harding J., & Balacheff N. (2006). Advances of semantic web for e-learning: expanding learning frontiers. *British Journal of Educational Technology*, 37 (3), pp. 321-330.
- Norman D.A. (1991). Cognitive artifacts. In J.M. Carroll (Eds.), *Designing Interaction: Psychology at the Human-Computer Interface*. Cambridge: University Press.
- Notti A.M. (Ed.), (2003). *Modelli statistici per la ricerca educativa*. Lecce: Pensa MultiMedia.
- Novak J.D., & Gowin D.B. (1989). *Imparando ad imparare*. Torino: SEI. (Edizione originale pubblicata 1984).
- Quillian M.R. (1968). Semantic Memory. In M. Minsky (Ed.), *Semantic Information Processing*. Cambridge: MIT Press (trad. it. *Un modello di memoria semantica*, VS, 1,1971).
- Rivoltella P.C. (2010). Rappresentare. Conoscenza, media, tecnologia. In A. Cattaneo, P.C. Rivoltella (Eds.), *Tecnologie, formazione, professioni. Idee e tecniche per l'innovazione*. Milano: UNICOPLI.
- Rossi P.G., & Toppano E. (2009). *Progettare nella società della conoscenza*. Roma: Carocci.
- Studer R., Benjamins V.R., & Fensel D. (1998). Knowledge engineering: principles and methods. *Data Knowledge Engineering*, 25, (1-2), pp. 161-197.
- Stutt A., & Motta E. (2004). Semantic Learning Webs. *Journal of Interactive Media in Education*, 10, pp. 253-285.
- Vygotskij L.S. (1974). *Storia dello sviluppo delle funzioni psichiche superiori e altri scritti*. Firenze: Giunti-Barbera (Edizione originale pubblicata 1931).

