Exploring attitudes of learners with respect to different learning strategies and performances using statistical methods

Silvia Rita Viola[1], Alberto Giretti[2] and Tommaso Leo[3]
[1] Ph. D. in "E-Learning", Dipartimento di Ingegneria Informatica, Gestionale e dell' Automazione, Università Politecnica delle Marche, Ancona, Italy
[2] Dipartimento di Architettura, Costruzioni e Strutture, Università Politecnica della Marche, Ancona, Italy
[3] Dipartimento di Ingegneria Informatica, Gestionale e dell'Automazione, Università Politecnica delle Marche, Ancona, Italy

***Abstract*— In this work the problem of identifying relationships between different learning strategies and learning outcomes is addressed.**

Classical statistical methods such as p values and chi square test, as well as Multiple Correspondence Analysis are employed; variables to be explained are performances of learners in Multiple Choice Tests (MCT) and Design Tests.

It is shown that: the methods are able to detect differences with a different sensitivity; the methods are able to detect characteristics belonging to the metacognitive domain; specific strategies are effective to learn complex skills.

Further applications are discussed, especially for what concerns cognitive and metacognitive changes happening in time.

***Index Terms*—E-Learning, Multivariate statistics, Learning process characterization, User models, Metacognition.**

I. INTRODUCTION

This work addresses the problem of identifying relationships between different cognitive and metacognitive strategies and learning outcomes of a set of students using navigation data in adaptively usable e-learning courses.

Data come from the WINDS Project, a fifth framework funded European Project that aims to support Architecture and Civil Engineering teaching for university students coming from different European countries. The pedagogical approach is inspired by a constructivist, interdisciplinary, intercultural and multidisciplinary view [4, 5].

Students use different kind of Los in the WINDS Advanced Learning Environment (ALE), including traditional objects, such that written texts and images, as well as multimedia and interactive objects, such that cases, or concepts or maps that allow student to personalize learning paths.

In this work some statistical analyses have been performfed in order to explore:

a) the statistical significance of the difference of the mean number of LOs used by each group of students;

b) the graphical and numerical exploration of differences between groups using Multiple Correspondence Analysis (MCA).

The approach follows a data-driven procedure [3] in which the model is built following a bottom-up inference.

Variables to be explored are:

a) performances of learners in multiple choice test;

b) performances of learners in design test.

Literature in fact suggests that different kinds of tests are related with different kinds of cognitive and metacognitive features of learning process. As an example, using Bloom's taxonomy of educational objectives [2], MCT is said to be able to assess Knowledge (1) and Comprehension (2) levels, while Design Test is said to be able to assess higher cognitive and metacognitive features, related to the levels going from Application (3) to Synthesis (5).

Moreover, it is to suppose, according to constructivist approaches to learning process, that many different strategies can be found looking at the usage of different kinds of objects; for what concerns this dataset, some of them have been empirically detected, namely traditional, active, collaborative, hypertextual strategies [10].

Such an analysis seems to be helpful in order to detect which strategies are more effective when the mediation of technology is introduced in the learning process, and what changes, from cognitive, metacognitive and pedagogical perspective, have to be related with the introduction of technologies into pedagogical settings; this point can be analyzed both from a shot-term view and from a long-term one.

It seems in fact that, when students starts to use WINDS ALE, they transfer strategies built in traditional settings in electronic environment, in which the effectiveness of strategies decreases. It could be then interesting to verify if some changes in strategies appear with the progression of the learning process and in which ways they impact on learning outcomes.

Literature gives different meanings to these hypotheses according to the different theories. In particular, the research on the impact of metacognition on learning process enlightens the importance of choosing the "best" strategies for an effective learning; moreover, the importance of changes of strategies according to contextualized settings and to the available artefacts and media is underlined.

From another viewpoint, Gregory Bateson describes learning phenomena in terms of "levels of learning". In [2] he indicates four levels of learning, numbered from zero to three, according to the different degree of generality of the changes implied in both the process of learning and in what has to be learned.

Inside Bateson's framework, levels zero and one refer, broadly speaking, to the learning of contents, while levels two and three refer to the learning of expertises and competences. In particular, learning two is characterized in terms of "learning to learn", that is the ability to learning *how* to learn, and level three is characterized in terms of learning to change the strategies acquired in the "learning two" step. Despite Bateson never talks about metacognition, his framework allow a general and flexilble characterization of learning from both cognitive and metacognitive viewpoint.

According to Bateson's framework, it seems reasonable to suppose that the transfer of strategies from traditional pedagogical settings to electronic ones reveal a difficulty related to "learning two" and "learning three" levels.

Eventually, from a methodological viewpoint, a preliminary comparison between the models obtained by Multiple Correspondence Analysis and Principal Component Analysis is provided. Principal Component Analysis models of learners profiles have shown their effectiveness in giving a meaningful characterization of learning process [10]. The results of comparison are shown despite it is not easy to give a clear interpretation, since many differences, for what concerns samples, methods and purposes are present in the two models.

II. MATERIALS AND PREPROCESSING

The analysis is based on a data set of students that was collected in experiments conduced with regular university European students within the WINDS project, a fifth FP IST funded project. The performances achieved by 34 students belonging to the dataset are considered inside the experimental setting. The frequencies of usage of each kind of LO and the performances achieved by these students are known.

The site of the Portal contains 22 courses at all; the sample was composed by a subset of students selected from students geographically distributed over Europe attending 8 Courses, namely:

1. Design and Computer Design Contemporary Architecture;
2. The Design of Network Space;
3. "La progettazione di spazi architettonici urbani nelle periferie interne della citta' storica", in italian (in english: "The project of architectonic urban spaces inside boundaries of hystorical town");
4. Construction Design Process;
5. Construction Quality, Safety and Environment Management;
6. Re-engineering the Construction Process;
7. Innovative materials and techniques for the environment sensible building envelope design;
8. Design of Structural Elements in Reinforced Concrete.

According to the inspiring criteria, WINDS ALE provides different kind of objects, devoted to promote an efficient learning in Design and Architecture.

Near to traditional Learning Objects containing lessons or self- evaluation tests, LOs supporting both active and collaborative learning are provided, according to the project inspiring criteria and guidelines. In particular, such kinds of objects are:

- "Cases", that are objects in which students are invited to analyze a real-world design task, realized by a famous practitioner, that is explained and commented in details; they are designed to support active learning; some of them uses multimedia objects;
- "Concepts", that are definitions of keywords occurring in paragraphs objects; both the number and the objects themselves change according to each selected paragraphs;
- "Maps", that are concepts maps provided by links accessible by concept pages, conceived to give a non linear and interdisciplinary view of each matter, by which learners can "jump" to other concepts, or paragraphs;
- "Annotations", that are a kind of "electronic notebooks" in which learners can put their observations; annotations can be viewed by anyone else and collaboratively edited and enriched;
- "Discussions", that are kind of forums accessible during navigation, and collaboratively enriched.

Figures 1 and 2 show respectively the ALE and a Concept Map.

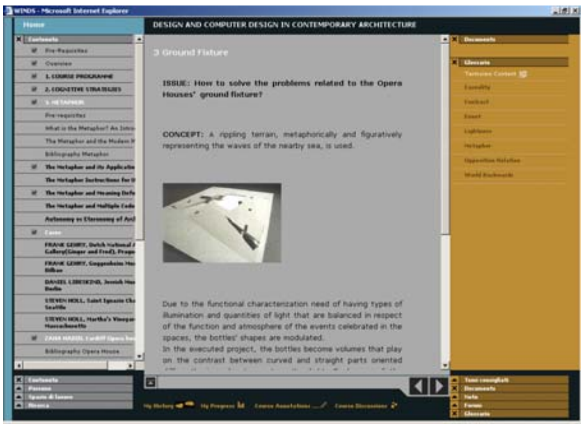

Figure 1. - The WINDS Advanced Learning Environment

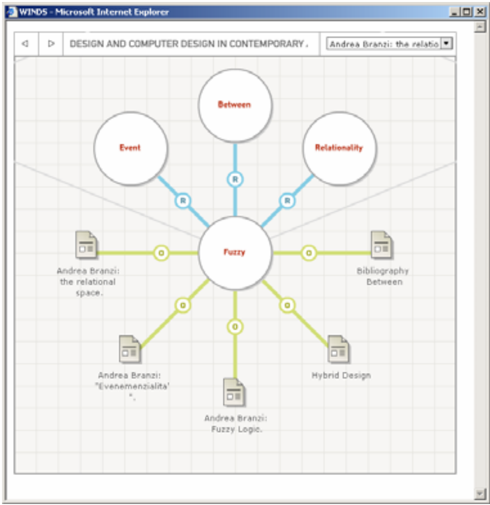

Figure 2. – A map

The topology of the Portal is made by “pointers” objects, such as Units, that contain links to other sections; Paragraphs; Concepts and Maps; a navigation tree and links to Annotations and Discussions are always available to the students. There are eventually “Courses” category pages - the Courses main pages - here considered as “dummy” pages sine they are an “obliged” way to reach other objects with zero degree of freedom.

WINDS ALE is provided by a database, devoted to keep tracks of all transactions; the database consists on more than 30 tables. In database each object is univocally labelled according to the categories above discussed.

For the purposes of analysis preprocessing steps consist in extracting all sessions for the entire sample of students; since every user was univocally identified by an ID number, sessions are univocally determined.

Data duplications and uninformative contents were removed.

Each transaction starting with a login action and ending with a logout action univocally associated with an user is considered a session; information about all visited pages, and time spent on each page are extracted.

Sessions where users see only one or more pointer page(s), namely Courses Homes or Units, without visiting anything else have been classified “dummy” and removed.

All pages contained in non dummy sessions are deserved and labelled according to the database categories, that is:

1. Units
2. Paragraphs
3. Cases
4. Exercises
5. Concepts
6. Annotations
7. Discussions
8. Maps

Data are then transferred into a matrix containing sessions in rows and frequencies in columns.

Each row contains the number of objects of each category viewed from a user or inside a session.

Scores are expressed in a numeral scale going from one to ten. The total score for each student is the average of the scores obtained by the student in different test sessions. Students have been then grouped according to the performances at tests, namely “high” ($\geq \frac{8}{10}$), “medium” ($\geq \frac{6}{10}, < \frac{8}{10}$) and “low” ($< \frac{6}{10}$).

Table 1 shows the size of each group together with the sample averages of MCT and Design scores.

TABLE I.
CARDINALITY OF GROUPS AND SAMPLE AVERAGE OF VOTES

Card. of groups	MCT	Design
High	28	20
Medium	3	7
Low	3	7
Sample average	8.51	7.78

III. METHODS

A. *Analysis of the significance of differences of the means*

The first analysis of the dataset was performed using classical statistical tests, such as *p* values and *chi square* test.

p values, as well as *chi square* test, are well known methods for hypothesis testing inside a classical experimental design setting; *p* values estimate the probability that the means of two or more samples are near to the value given by the estimate. If values of *p* are small, then the null hypothesis is rejected. The level of significance is set at $\alpha = .05$.

Chi square test is devoted to test independence comparing the differences between expected values and observed ones.

Since the dataset available for the analysis was not collected according to any experimental design setting, no assumption can be made about the underlying distribution.

B. *Multiple Correspondence Analysis*

Multiple Correspondence Analysis (MCA) is a well-known multivariate statistical method for data dimensionality reduction and graphical exploration of data, especially for what concerns categorical data and non homogeneous data [6, 8, 9].

MCA performs the optimal projection of a matrix in rows and columns space simultaneously, using *chi square* metric.

MCA considers the bipartite graph $G=(V,E)$, $V=\{V_1,V_2\}$ whose set of vertices are the individuals (V_1) and variables (V_2).

Broadly speaking, MCA performs the Singular Value Decomposition (SVD) of the Adjacency Matrix *A* of the graph *G*, looking for the independence of each value computed according to *chi square* metric given by

$$\chi^2_{i,j}=\sum_i\sum_j\frac{\left(a_{ij}-E(a_{ij})\right)^2}{E(a_{ij})} \tag{1}$$

where $E(a_{ij})$ is the expected value of the entry *i,j* of the matrix *A*.

The aim of the analysis is to search for values far from independence, that indicate an association, both positive and negative, between individuals and variables. The association is detected using the different numerical relevance when independence condition does not hold.

Many algorithms and procedures are given in order to perform MCA: see [6, 8, 9].

In this study Greenacre's procedure is preferred, being the generalization of the Theorem of the best *k*-rank approximation [6] of a matrix obtained by the SVD to matrices weighted by diagonal positive definite matrices.

The matrix

$$\hat{A}=\frac{1}{N}A, A\in\mathbb{R}^{mxn}, N=\sum_i\sum_j a_{ij} \tag{2}$$

is considered, being *A* the adjacency matrix of the graph containing in rows students and in columns frequencies of objects.

Let

$$r=\hat{A}i_c, i_c\in\mathbb{R}^{nx1}, i_c=[1,...,1]' \tag{3}$$

$$c=\hat{A}'i_r, i_r\in\mathbb{R}^{mx1}, i_r=[1,...,1]' \tag{4}$$

be the vectors containing the marginal frequencies of rows and columns of the matrix *A*, such that, being verified the hypothesis of complete independence, the matrix $\hat{A}$ is given by

$$\hat{A}=Nrc'. \tag{5}$$

The matrix

$$.\hat{A}_{res}=\hat{A}-rc' \tag{6}$$

containing residuals from independence according to *chi square* metric is weighted, both in rows and in columns space, and decomposed as follows

$$.\tilde{A}=\Omega^{1/2}\hat{A}_{res}\Phi^{1/2}, \tag{7}$$

where

$$\Omega=diag(i_r), \Omega\in\mathbb{R}^{mxm}, \tag{8}$$

$$\Phi=diag(i_c), \Phi\in\mathbb{R}^{nxn} \tag{9}$$

are the diagonal matrices containing the square roots of the marginal frequencies in rows and in columns space.

From a geometric viewpoint, MCA realizes an orthogonal transformation that emphasizes the polarization of elements far from independence, both for positive values and for negative values, in which distances, calculated according to *chi square* metric, represent the deviance from independence.

Linear combinations to be considered are then

$$Y_r=\tilde{A}\tilde{V} \text{ in row space} \tag{10}$$

$$Y_c=\tilde{U}'\tilde{A} \text{ in column space.} \tag{11}$$

IV. RESULTS

A. *Multiple Choice Test Performances*

Results show that, on average, the same set of students achieves a greater average performance in MCT than in design test; the distribution of the students inside groups confirms this result. This result could be related both with the different level of difficulty of the two tests, confirmed by Bloom's taxonomy, and with the different strategies used in the two cases and revealed by the analysis. Moreover, the consistent difference of cardinality of the groups requires adequate tests in order to avoid biased results: in particular, an underlying gaussian distribution can not be assumed.

Looking at intra-groups averages, it is evident that low performances (LP) profiles show a scarce usage of almost all electronic resources, both from a qualitative and from a quantitative viewpoint: the most used categories of objects are in fact traditional objects; a very limited usage of cases ("active" objects) is done; no one of the other objects is used.

High performances (HP) and medium performances (MP) profiles show instead more variability; moreover, interesting relationships hold between LP and MP. HP profiles share in fact an equilibrium, with a limited between-groups variability, in traditional objects usage with MP ones.

HP profiles show
a) more variability in the usage of all kinds of objects;
b) a greater usage of objects made available by technologies-mediated environments, such as linked resources, hypermedia, objects stimulating active and collaborative learning.

In particular, the frequency usage of cases and exercises, that encourage active learning strategies, is double in MP profiles: this is related with performances, since active strategies of learning affects only superficially levels one and two of Bloom's taxonomy, and then it seems appropriate to talk about inadequate strategies with respect to objectives.

The analysis of intra-groups correlations shows that for HP group a high correlation (greater than .7) is given by collaborative objects combinations, despite is not easy to associate a precise meaning to this mark.

Tables 2 and 3 show the means of the three groups and the statistical significance at the level of $\alpha = 0.05$ of both of *p* values and of *chi square* test.

Both *p* values and chi square values are far from significance.

TABLE II.
MCT STATISTICAL SIGNIFICANCE

Parameters	Values
p	.436
chi-square	16.2451
d.o.f.	16
significance .05	no

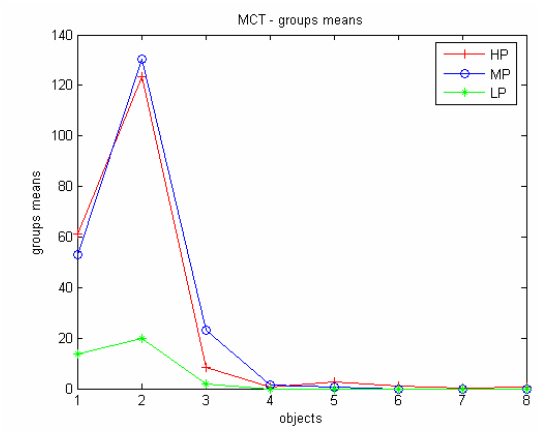

Figure 3. - Plot of the means of the three different groups with respect to MCT tests performances. Objects are: 1. units; 2. paragraphs; 3. cases; 4. exercises; 5 concepts; 6. annotations; 7. discussions; 8. maps.

TABLE III.
MEANS OF THE THREE DIFFERENT GROUPS WITH RESPECT TO MCT TESTS PERFOMANCES

Objects	High	Medium	Low
Units	61.107	53	13.667
Paragraphs	123.36	130.33	20
Cases	8.5714	23	1.6667
Exercises	0.67857	1.3333	0
Concepts	2.6071	0.66667	0
Annotations	0.92857	0	0
Discussions	0.21429	0	0
Maps	0.46429	0	0
Size	28	3	3

In order to cross-validate the results, as well as for graphical exploration of relationships, MCA was used.

MCA detects different models belonging to the three groups. In particular, differences could be appreciated in terms of:

a. sign of each component of the RSVs vectors;
b. absolute values of the components.

Table 4 shows the comparison of the second and third axes of MCA models respectively of HP, MP and LP groups. Axes refers to the right singular vectors (RVSs) obtained by the SVD, being the largest singular value of the weighted matrix SV decomposed $\sigma_1(\tilde{A}) < 1$ [6, 8].

HP group axes show the greatest, and most significant, combinations, due both to the number of objects employed and to the frequencies of usage of each object.

For what concerns HP, objects show in fact coherence with traditional (units, paragraphs), hypertextual and non sequential (concepts, maps), collaborative and active (cases) clusters of variables; for what concerns MP, the opposite sign on the third axis of paragraphs with respect to cases seems to be related with the opposition of active/traditional strategies usage.

This significance of combinations is also expressed graphically by MCA geometry: MCA geometry clusters (figures 6, 7 and 8) show to be significant from a cognitive and metacognitive viewpoint.

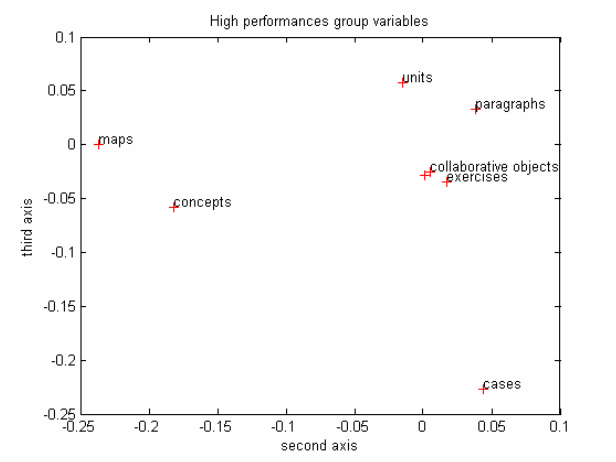

Figure 4. – MCA geometry of Multiple Choice Tests High Profiles

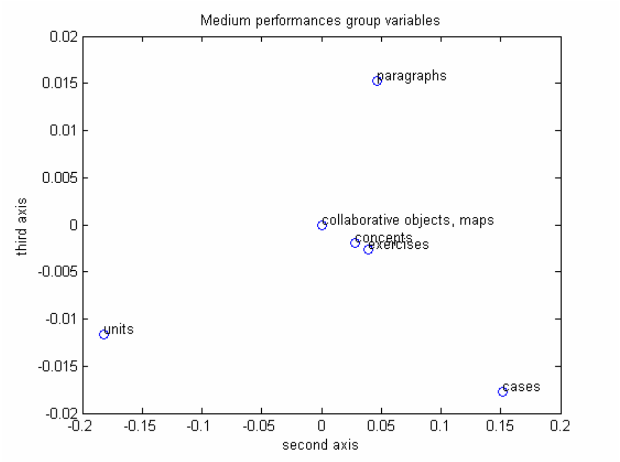

Figure 5. – MCA geometry of Multiple Choice Tests Medium Profiles

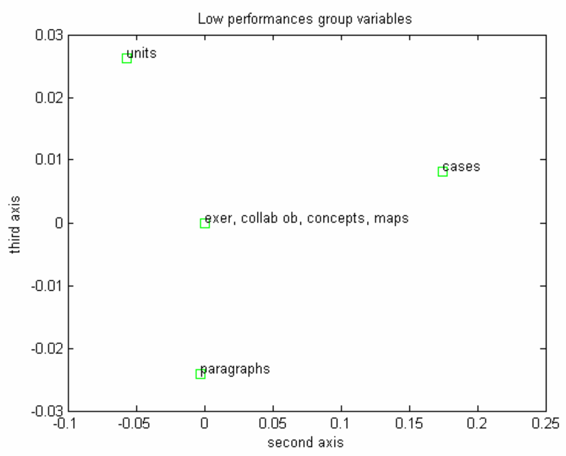

Figure 6. – MCA geometry of Multiple Choice Tests Low Profiles

TABLE IV.
RSV OF MCT MODELS BELONGING TO THE THREE GROUPS AND COMPARISON WITH 57 INDIVIDUALS SAMPLE PCA

Objects	High	Medium	Low	PCA(1)
SECOND AXIS				
units	-0.04867	-0.74002	-0.31125	0.890
paragraphs	0.12484	0.18932	-0.017004	-0.379
cases	0.14476	0.61479	0.95018	-0.239
exercises	0.058172	0.16032	0	-0.017
concepts	-0.59502	0.11337	0	0.055
annotations	0.004707	0	0	-0.028
discussions	0.01662	0	0	-0.004
maps	-0.77676	0	0	0.033
THIRD AXIS				
units	0.23275	-0.44288	0.71856	0.185
paragraphs	0.13451	0.58014	-0.65854	-0.170
cases	-0.91163	-0.67241	0.22359	0.966
exercises	-0.13821	-0.10055	0	0.006
concepts	-0.23339	-0.07109	0	0.049
annotations	-0.11226	0	0	0.014
discussions	-0.10234	0	0	0.006
maps	0.002698	0	0	-0.015

(1) This column refers to the second and third RSVs obtained by the SVD of the matrix containing learner profiles in rows and frequencies of objects in columns described in [10].

B. *Design Test Performances*

Design test results show more equilibrium in the cardinality of the three groups.

LP profiles distinguish again themselves from the others groups for the scarce usage of resources.

For what concerns HP and MP profiles, the limited between-groups variability in using traditional objects is detected, as in the case of MCT; moreover, between-groups variability is given by active objects (cases, exercises) usage: in fact HP profiles show an average usage of cases and exercises that is approximately two times the one detected in MP; it seems that no significant difference is present for what concerns traditional objects, concepts and discussions usage, while the number of annotations and maps used decreases.

For what concerns MP, both concepts and collaborative objects show relevant correlations (> .8) with other kinds of objects, while maps do not.

Tables 5 and 6 show the means of the three groups and the statistical significance at the level of $\alpha = 0.05$ of both of *p* values and of *chi square* test.

Both *p* values and *chi square* values are significant. In particular, *p* values are of an order of magnitude smaller than the ones obtained for MCT tests.

A further analysis shows that the significance could be related both to groups structure, that are well separated in terms of characterizing dimensions, and to the greater equilibrium shown by groups cardinalities.

TABLE V.
MEANS OF THE THREE DIFFERENT GROUPS WITH RESPECT TO DESIGN TESTS PERFOMANCES

Objects	High	Medium	Low
Units	49.85	68.857	61.714
Paragraphs	107.8	142.71	107.14
Cases	13.6	6	0
Exercises	1	0.42857	0
Concepts	2.7	2.8571	0.14286
Annotations	0.65	1.8571	0
Discussions	0.2	0.28571	0
Maps	0.35	0.85714	0
Size	20	7	7

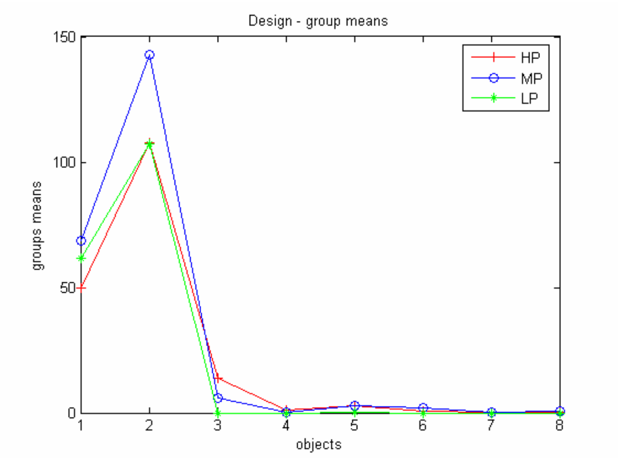

Figure 7. – Plot of the means of the three different groups with respect to Design tests performances. Objects are: 1. units; 2. paragraphs; 3. cases; 4. exercises; 5 concepts; 6. annotations; 7. discussions; 8. maps.

TABLE VI.
DESIGN STATISTICAL SIGNIFICANCE

Parameters	Values
p	.0436
chi-square	26.8179
d.o.f.	16
significance .05	yes

Looking at differences for what concerns design test, that evaluates more complex skills (such as the ones going from level 3 lo level 6 of Bloom's taxonomy) is evident the different distribution of the number of objects. Here, cases and exercises are positively correlated with level of performances; moreover, for what concerns units and paragraphs, the performances seems to be not dependent from the number of objects requested.

It could be argued that these results depend on metacognitive characteristics of learners more than on cognitive ones, being the process of learning complex skills different.

This seems also to confirm the relative independence of performances from the number of resources used.

Moreover, the importance of active and constructivist ways of learning for acquiring complex skills is evident looking at results, especially for what concerns the usage of cases and exercises.

It seems instead that non sequential objects, like concepts and maps provided by the system, do not affect deeply the outcomes.

MCA geometry also shows coherence with cognitive unobservable dimensions.

In particular, clusters related with cognitive and metacognitive strategies are detected in HP profiles model, as for MCT HP profiles; for what concerns MP, the model does not detect clearly clusters coherent with traditional, collaborative, non sequential strategies; moreover the MP high correlations (> .8) seems to be due to non linear relationships between variables, while cases and units are the only well separated variables.

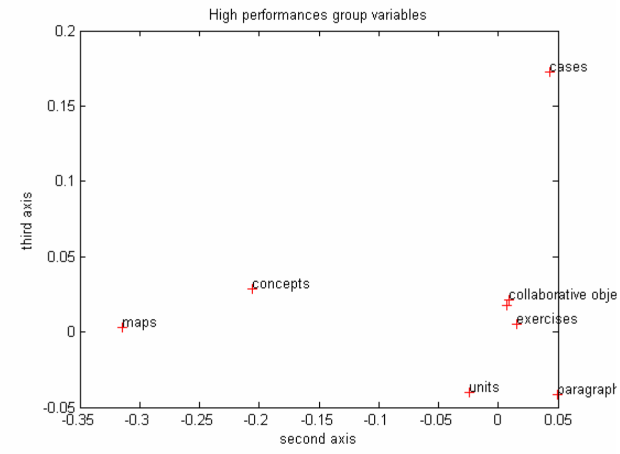

Figure 8. – MCA geometry of Design Tests High Profiles

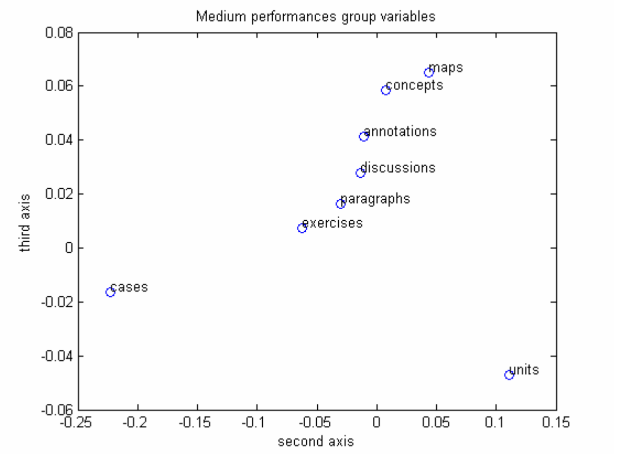

Figure 9. – MCA geometry of Design Tests Medium Profiles

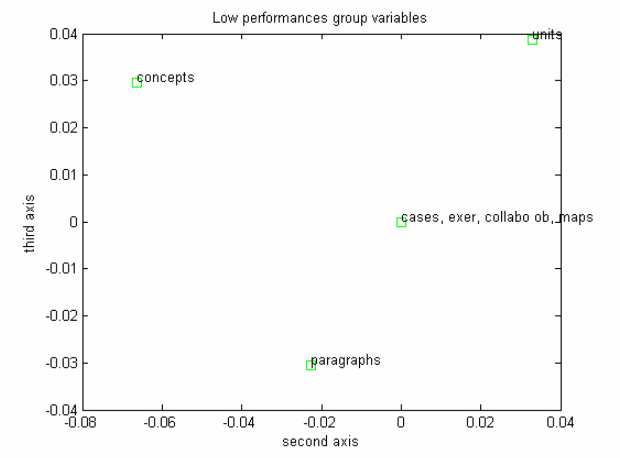

Figure 10. – MCA geometry of Design Tests Low Profiles

TABLE VII.
RSV OF DESIGN MODELS BELONGING TO THE THREE GROUPS AND COMPARISON WITH 57 INDIVIDUALS SAMPLE PCA

Objects	High	Medium	Low	PCA(1)
SECOND AXIS				
units	-0.06402	0.42059	0.42548	0.890
paragraphs	0.12851	-0.11476	-0.29163	-0.369
cases	0.11334	-0.84879	0	-0.239
exercises	0.041353	-0.23939	0	-0.017
concepts	-0.53748	0.02741	-0.85669	0.055
annotations	0.023843	-0.04068	0	-0.028
discussions	0.018853	-0.05046	0	-0.004
maps	-0.8216	0.16503	0	0.033
THIRD AXIS				
units	-0.21699	-0.41286	0.67364	0.185
paragraphs	-0.2221	-0.14568	0	-0.170
cases	0.92591	-0.14568	0	0.966
exercises	0.002782	0.006329	0	0.006
concepts	0.15401	0.51357	0.51501	0.049
annotations	0.11374	0.36218	0	0.014
discussions	0.092744	0.24441	0	0.006
maps	0.015982	0.57347	0	-0.015

(1) This column refers to the second and third RSVs obtained by the SVD of the matrix containing learner profiles in rows and frequencies of objects in columns described in [10].

For what concerns meaningfulness of learning process characterization, results seems to confirm the ones obtained in a previous work [10], taking into account both differences of samples and methodological differences between MCA and PCA [7] (tables 4 and 7, “57 ind. Sample PCA” column, show the same axes of the whole sample model obtained by PCA). In particular, the same latent dimensions (i.e. traditional, active, collaborative, hypertextual) are detected, despite a different representation is given according to each model.

V. Conclusions

From a methodological viewpoint, it is shown that multivariate methods could be more sensitive than traditional statistical tests in detecting differences outside from any experimental design setting. Moreover, a customized approach, that satisfy the need for a more accurate modelling of groups specificities, seems to be more appropriate in order to detect and explain differences: the comparison of different models, each of them built on a specific group, seems to detect differences better than the comparison of different groups inside a unique model.

A comparison with the 57 individuals sample PCA [10] second and third axes show a different distribution of variance.

The preliminary results show that the models obtained using the two methods are provided by a different meaningfulness in learning process characterization. These preliminary results confirm the ones given by literature, despite further analyses focused on the behaviour of the methods are needed.

For what concerns cognitive and metacognitive viewpoints, it seems that the usage of different kinds of objects, such that traditional objects or web enhanced objects, could reveal some of the metacognitive features of learners together with cognitive ones.

In particular, when various LOs and various tasks, such as project tasks together with traditional tasks, are provided by a learning environment, it is possible to infer information on metacognitive strategies of learners looking at usage data together with the outcome of assessment of learners themselves.

From these results it seems reasonable to relate outcomes of learning to *quality* of cognitive and metacognitive strategies rather than to *quantity* of objects viewed.

Moreover, the hypothesis of a difficulty in “learning to learn” and “tertiary learning”, that is, learning to modify what has been learned (Gregory Bateson, [2]) seems to be likely when the electronic media are employed in educational contexts. It has been in fact shown that low performances are related with traditional strategies, i.e. strategies that do not exploit the potential of technologies-mediated learning.

High performances seems to be the outcome of a complex metacognitive structure of some learners that are able to develop the skill of learning new cognitive strategies in order to adapt better to the electronic environment.

In order to monitor the evolution of the learning process along time, the methods here applied seem to be effective and promising, but they have to be adapted to changing situations, by improving their flexibility.

In particular, it seems that appropriate distance or proximity measures able to characterize profiles, as well as changes of profiles in time, could greatly improve accuracy and meaningfulness of these models.

Further application domains can also be considered, especially in classification and monitoring systems that are focused on the learning process.

Eventually, the need for tracking metacognitive changes along time became more and more present in E-Learning systems, and these methods have shown to be effective in revealing trends and changes happening in time.

ACKNOWLEDGMENT

S.R.V. thanks Gianna Maria Del Corso and Anna Maria Perdon for their hepful suggestions and their warm encouragement.

REFERENCES

[1] G. Bateson, *Steps to an Ecology of Mind*, 1st ed., San Francisco: Chandler Publishing Company, 1972.

[2] B. S. Bloom, M. D. Engelmann, E. J. Furst, W. H. Hill, D. R. Krathwol, *Taxonomy of Educational Objectives: the Classification of Educational Goals, Handbook 1: Cognitive Domain*, New York: McKay Publishing, 1956.

[3] V. Cherkassky, F. Mulier, *Learning from data*, New York: Wiley, 1998.

[4] M. De Grassi, *Web Based Intelligent Tutoring Design System, Final Report, Part I, Overwiev*, WINDS IST 1999 10253, 1999.

[5] M. De Grassi, *Web Based Intelligent Tutoring Design System, Final Report, Part II, Methodologies*, WINDS IST 1999 10253, 1999.

[6] M. Greenacre, Theory *and Applications of Correspondence Analysis*, London: Academic Press, 1984.

[7] T. I. Jolliffe, *Principal Component Analysis*, New York: Springer, 1986.

[8] L. Lebart, A. Morineau, K. M. Warwick, *Multivariate Descriptive Statistical Analysis*, New York: Wiley, 1984.

[9] J. Michailidis, J. De Leeuw, "The Gifi system for Descriptive Multivariate Analysis", *Statistical Science*, vol. 13, pp.307-336, 1998

[10] S. R. Viola, A. Giretti, T. Leo, "Discovering learning process patterns by multivariate analysis of usage frequencies data in e-learning courses", *ICL 2005 Proceedings*, Kassel University Press, ISBN 3-89958-136-9, 2005

AUTHORS

Silvia Rita Viola, born in 1976, is Ph. D. in E-Learning. She is with the Università Politecnica delle Marche, 60100 Ancona, Italy (e-mail: sr.viola@gmail.com).

Alberto Giretti, is Research Scientist at the Dipartimento di Architettura, Costruzioni e Strutture (DACS), Universita' Politecnica delle Marche, 60100 Ancona, Italy (e-mail: a.giretti@univpm.it).

Tommaso Leo, born in 1944, is since 1981 with the Dipartimento di Ingegneria Informatica, Gestionale e dell' Automazione (DIIGA), Università Politecnica delle Marche, 60100 Ancona, as Full Professor at the Chair of Automatic Control (e-mail: tommaso.leo@univpm.it).

Manuscript received 23 April 2006.