EFFECTIVENESS OF A HYPERMEDIA TOOL IN THE INNOVATIVE TEACHING OF PNEUMATIC ENGINEERING DESIGN DRAWING

Ramos Barbero B.; Pelaez Vara J.; García Maté E.; Ruiz Calvo J.
Área de Expresión Gráfica en la Ingeniería.
Universidad de Burgos (España)
Escuela Politécnica Superior.
Avda. General Vigón s/n 09006 Burgos. (España)
E-mail: bramos@ubu.es
Tel: 00 34 947-258926 / 258927

ABSTRACT

This article presents research undertaken to determine the quality, effectiveness and efficiency of the "pneumatic engineering design drawing" hypermedia. Among our reasons for carrying out this research are the implementation of the ECTS (European Credit Transfer System), the inexistence on the market of a similar product, and an increasing daily need for asynchronous teaching methods. Its objectives sought to evaluate the application and its impact on the teaching–learning of pneumatics. Two sources of documentation were used as research strategies: opinion polls and experimentation. Three groups of equal size were formed at random, based on traditional teaching (T), self-study using multimedia (M), and traditional teaching reinforced by the use of the hypermedia (A). The following rank among some of the more noteworthy conclusions. The hypermedia application has been positively rated. In the group with prior knowledge of the subject, the results of the post-test are strongly related to those of the pre-test in pneumatics. There is no empirical evidence to support the affirmation that students in group A achieved better results that those in group M or group T. In group M, academic performance is not directly dependent on the intellectual abilities of the student. The minimum time-saving in group M was 42% in comparison with the other two groups.

Keywords: multimedia, hypermedia, computer-aided instruction, science education, training, learning, effectiveness, efficiency, achievement assessment.

1. INTRODUCTION

The "pneumatic engineering design drawing" hypermedia was created to improve the process of teaching/learning pneumatics. This improvement can be brought about by increasing the effectiveness of the process: increasing the pass rate, the degree to which knowledge is acquired, enabling learning in places other than the lecture hall, and providing students with a system that allows them to monitor progress in acquiring knowledge. Alternatively, improvements to the efficiency of the process can be brought about, either due to a reduction in lecturing hours, by centring the classes on those concepts that present the greatest difficulty to the student; or, due to a reduction in the time needed to acquire the proposed didactic objectives.

This research is intended to reveal the quality of the tool that has been developed and its impact on the results of the learning process, measuring those results in terms of the knowledge acquired, its influence on different types of students, and the time spent studying the subject matter.

To validate the research results, the effect of disturbing variables had to be neutralised, such as intellectual aptitude, prior knowledge of pneumatics, or the experimental conditions. The reliability of the questionnaires and the design of the tests also had to be carefully checked.

The results of the research provide valuable information for the implementation of the ECTS system, since they provide guarantees as to the quality and effectiveness of the tool, as well as a good estimate of the time needed, by both the lecturer, and the student, to develop and to study the subject matter.

2. FORMULATION OF THE RESEARCH PROBLEM

The reasons that have led us to develop the "pneumatic engineering design drawing" hypermedia are as follows:

- The need to implement, before 2010, the new curriculum based on the ECTS (European Credit Transfer System) system. This system calls for less lecture hall presentations, a greater student workload and intensification of the tutorial programme by the university lecturer. This means that systems have to be created to help the student study the subject matter independently, to help the lecturer to identify the concepts that cause greatest difficulty for students, and to facilitate the development of tutorials.
- There is no pneumatics-based application available on the market that contains a navigation system that serves as a private study guide to the subject matter for students and monitors their progress in learning the subject matter at any given moment, as well as monitoring those areas that need to be covered in greater depth.
- The heterogeneity of students in terms of intellectual aptitude, availability throughout the day and lodgings, mean that there is an ever-increasing need for asynchronous teaching methods, in which students impose their own pace of study, at a time and place that they themselves decide.

Following the creation of the pneumatic design hypermedia, the following research-related questions arose: how well does the hypermedia convey the information that it contains? Do its contents fit in well with the proposed didactic objectives? Do students who use the hypermedia achieve better results? What sort of student responds best to the use of this tool? Is there any relationship between study time and academic achievement? In order to respond to these questions, the objectives, we set ourselves the following objectives:

1. To evaluate the programme as a means of conveying information
2. To evaluate the contents of the programme
3. To analyze the influence of the hypermedia on learning outcomes in the field of pneumatics by carrying out research involving 2nd year students studying for a degree in industrial engineering at the University of Burgos.

The reasoning set out above, the ensuing questions, and the evident interest of pneumatics for the industrial engineer, all justify the proposed objectives; in order for lectures to focus on those concepts that pose the greatest difficulty to students; so that the didactic objectives may be achieved within the intervals of time set out in the new study plans.

3. APPLICATION OF OPERATIONAL RESEARCH TO THE PROBLEM

Having formulated the questions and research objectives, operational research is then applied to the problem [1,2,3,4]. Operational research encompasses two concepts: conceptualisation and measurement.

Conceptualisation consists in extracting the questions and objectives of the research, and a series of hypotheses constituting probable outcomes. These are put forward by the researcher in response to the research questions and objectives. In our case, the following hypotheses are advanced:

1. The pneumatic engineering design hypermedia as a medium through which to convey information is of an acceptable quality.
2. In relation to the proposed didactic objectives, the contents developed in the pneumatic engineering design hypermedia are acceptable.
3. Significant differences are found between the scores obtained by students prior to studying and after having studied the subject matter.
4. Those students who attend lectures and who furthermore have access to the hypermedia application perform better than those who only attend lectures and those who only use the hypermedia.
5. The influence of the hypermedia is greater in those students with less intellectual aptitude than in those students with greater intellectual aptitude.
6. Regardless of any prior knowledge and the capacity of the student, in the self-study group, the academic performance is unrelated to the study time spent on the hypermedia application.

These hypotheses refer to concepts that indicate what has to be analysed. To do so, the hypotheses have to be translated into empirical variables or indicators, which will allow us to measure the concepts to be studied.

Measurement consists in assigning values or categories, depending on the type of variable, according to a set of pre-determined rules, to the different indicators that constitute the object under study.

4. THE RESEARCH DESIGN

The following strategies guided the research into the pneumatic engineering design hypermedia:

- reference to existing academic literature and statistics
- surveys
- experimentation

Whereas reference to academic literature and statistics constitutes a basic strategy that allows greater familiarisation with the research and guides us in the design and subsequent analysis of the results, the survey determines the quality of the hypermedia, and experimentation determines the effectiveness of this tool with regard to the learning of pneumatics.

4.1. EXISTING ACADEMIC LITERATURE AND STATISTICS

As regards the existing academic literature and statistics, the studies undertaken are very varied. This variety is not only due to the material being studied, as there are also other factors, such as the population profile, the proposed objectives of the research, the design of the research, as well as the period in which it is to be performed, which will influence the results of the research.

As regards the population, it might be academic [4,5,6,7] or not academic [2,3], better results being obtained from academic populations, as well as greater satisfaction with the use of these tools, in comparison with traditional teaching methods in a lecture hall.

As regards the objectives, whereas [2] is interested in establishing whether multimedia can be used to impart a course of lessons rapidly to users from different cultures and languages; [3,4,5,6] wish to establish the effectiveness of these tools, in the process of learning the subject matter; [7] expresses an interest in understanding the influence of the relationship between cognitive and demographic variables on the results of learning; and [8] establishes that students need support mechanisms and identifies the areas of the subject matter in which it is required.

Regarding the design of the research, in [2,6,7], no survey is undertaken, for which reason any poor results might be due to the subject matter being poorly structured, developed or presented, and not to the use of the tool. In [2,4], no pre-test is used to measure prior knowledge, which calls for great familiarity with the past academic experience of the student, to be sure that the results obtained are the direct consequence of using the tool. Another source of variability in the design of the research is that found by Jonassen (1994), who coincides with Drisoll and Dick (1999), in indicating that the majority of studies undertaken were carried out on populations for whom the subject matter lacked any interest [3].

Taking account of the period during which the research was carried out, in a meta-analysis of 248 studies on the effectiveness of Computer-Assisted Instruction (CAI), carried out between 1974 and 1986, Kulik and Kulik (1991) [3] concluded that the same tool exercised greater influence than suggested in a similar study undertaken by the same authors between 1966 and 1974. Along similar lines, Fletcher-Flinn and Gravatt (1995) carried out a meta-analysis between 1987 and 1992, into the influence of CAI compared with other teaching methods, observing that over recent years the tool exercised greater influence.

Other authors such as Niemiec and Walberg (1987), Schemeeckle (1998), and Dillon and Gabbard (1998), also observed modest increases in the results related to CAI compared with traditional teaching methods [3].

Authors such as Skinner (1990) [3], and Peláez (2001) [4], studied the effect of CAI, when used in conjunction with traditional methods and not only as a substitute. They concluded that the results of those students who, apart from attending lectures, had access to CAI, were better than those who were taught using more traditional methods.

4.2. SAMPLING

In order to comply with research objectives and test their hypotheses, the necessary samples are taken from the group to allow us to extract the desired information. These results, from which inferences can be made regarding the original survey population and the internal validity of the study, will depend on the sample size and the way in which it is taken.

In our case, the three following groups were formed from the group of 151 students who voluntarily participated in the study, of whom 87 were specialising in mechanical engineering and 64 in electrical engineering:

- Traditional group or control group, referred to as (T). This group was made up of 50 students who attended formal lectures, in which the whiteboard and transparencies were the most common support mediums used by the university lecturer.
- Support Group, termed (A). This group was made up of 50 students attending formal lectures but, in this case, the same university lecturer replaced the whiteboard and transparencies with the hypermedia as a support medium during the lecture and, in addition, students had access to the same tool to study pneumatics in their own time.
- Multimedia group, referred to as (M). This group was made up of 51 students, who did not attend lectures and who were able to access the hypermedia, as an additional tool, in their programme of independent, private study.

With the aim of ensuring internal validity in the study, the cognitive skills of the students were measured and they took a pre-test, in order to measure any prior knowledge they might have on pneumatics.

The Department of Psychology of the University of Burgos was consulted on the question of measuring cognitive skills in pneumatic design and drawing. Its recommendation was that for tasks related to visualization, interpretation and manipulation of objects, the most appropriate tests are spatial reasoning SR [9] and mechanical reasoning MR [10]. The validity of these tests is widely endorsed by a great number of psychologists.

To ensure content validity in the tests on pneumatics, 12 questions were prepared, for each of the proposed didactic objectives, from which 4 questions per objective for the pre-test, and another 4 questions per objective for the post-test , were all selected at random.

Having performed the SR and MR tests, the average of both tests f(SR, MR) was calculated, and the students were broken down into different strata as shown in table 1, in such a way that students at the same level were both equitably and randomly distributed across the three previously mentioned groups. By doing so, an effort was made to neutralise the influence that the disturbing variables might have on the relationship between teaching mediums and academic performance obtained by students in the post-test.

Strata	Pre-test = x	f(SR, MR) = y	Nos students	Nos students/group
SR, MR Test	Students with no prior knowledge	y ≥ 70	63	21
		50 ≤ y < 70	27	9
		y < 50	39	13
Pre-test Students with prior knowledge	x ≥ 7	y ≥ 70	3	1
	5 ≤ x < 7	y ≥ 70	4	1
	x < 5	y ≥ 70	3	1
	x ≥ 7	50 ≤ y < 70	0	0
	5 ≤ x < 7	50 ≤ y < 70	4	1
	x < 5	50 ≤ y < 70	0	0
	x ≥ 7	y < 50	3	1
	5 ≤ x < 7	y < 50	2	1
	x < 5	y < 50	3	1

Table 1. Population stratification

The random distribution of the students in different stratum gave rise to the results shown in tables 2 and 3.

TOTAL N^{O}. OF STUDENTS	TOTAL N^{O}. OF STUDENTS WITH PRIOR KNOWLEDGE	BRANCH (SPECIALISM)		GROUPS		SUB-GROUPS TREATED	
TOTAL Nos.=151 MEAN=1.3	Nº=22 (14.5%) MEAN=6.1	MECHANICAL Nº TOTAL=87	Nº=12(13.8%) MEAN=6.5	(TUESDAY) Nº TOTAL=71	Nº=11 (15.5%) MEAN=6.3	**A** TOTAL Nos.= 50	Nº=7 (14%) MEAN=6.2
						M TOTAL Nos.= 51	Nº=8 (15.6%) MEAN=6
		ELECTRICAL Nº TOTAL=66	Nº=10(15.1%) MEAN=5.5	(WEDNESDAY) Nº TOTAL=79	Nº=11 (13.9%) MEAN=5.9	**T** TOTAL Nos.= 50	Nº=7 (14%) MEAN=6

Table 2. Distribution of students with prior knowledge

Total Nos. Students	Specialism		Groups by specialism		Groups treated	
Nº = 151 Mean =63.1 DEV=24.8	MECHANICAL	Nº = 87 Mean=64.4 DEV =26.4	A	Nº = 29 Mean=65.2 DEV =26	A	Nº =50 Mean=63 DEV =24.2
			M	Nº = 29 Mean=63.5 DEV =27.7		
			T	Nº = 29 Mean=64.4 DEV =26.6	M	Nº =51 Mean=62.5 DEV =25.8
	ELECTRICAL	Nº = 64 Mean=61.3 DEV =22.3	A	Nº = 21 Mean=59.9 DEV =21.6		
			M	Nº = 22 Mean=61.1 DEV =23.3	T	Nº =50 Mean=63.9 DEV =24.9
			T	Nº = 21 Mean=63.5 DEV =23		

Table 3. Global distribution of students

As the sample was arrived at to obtain three groups of equal size, it must be asked whether it is representative of the group as a whole. To do so, we had recourse to the expression (1) taken from [1], that gives us the sampling error for a proportional stratified random sample.

$$E(p) = Z \cdot \sqrt{\frac{\sum_{i=1}^{n} n_i \cdot P_i \cdot Q_i}{n^2}} \qquad (1)$$

in which:

- E(p) is the sampling error, which represents the divergence between the population parameter and the estimate of the proportion.
- Z is the number of typical deviation units. Selecting Z = 2 gives a probability of success of 95.5% in the estimate of the population parameter.
- n_i is the number of units of each stratum that are found in the sample, table 1.
- P_i is the probability of exceeding the post-test in each stratum and Q_i is the probability of it not being exceeded. The supposition in our case is the least favourable in which $P_i = Q_i = 50\%$.
- n is the total size of the sample

By applying (1), we obtain E(p) = ± 14%. What this tells us is that in the least favourable case, there is a 95.5% probability that the percentage of passes in the post-test will lie close to the estimated value of ± 14%.

4.3. SURVEY AND EXPERIMENTATION

Two surveys ranking quality were used in order to measure the utility of the hypermedia,. One measured the underlying quality of the information presented, and another the underlying quality of the contents and their development. The two questionnaires contained questions that have been used in other research work [4,11], which were adapted, in some cases, for our work.

To collect the data on experimentation, the post-test was used. In the evaluation of results, statistical techniques were used with a level of significance of 5%, which gave a confidence level of 95%, when mapping the results taken from the sample onto the population parameters.

For the pre-test and post-test scores, the expression in (2) was used, as proposed in [12], for questions in which errors are penalised.

$$P = \frac{A - \frac{E}{K-1}}{N} \cdot 10 \qquad (2)$$

In which: P = test score, A = number of correct responses, E = number of erroneous responses, K = number of alternative responses to the questions, and N = number of test questions.

In order to measure the reliability of the pre-test and the post-test, the split-halves method was used [13], and the reliability for each student was determined by using the Rulón formula (3)

$$r_{xx} = 1 - \frac{Sp^2 - Si^2}{Sx^2} \qquad (3)$$

in which, Sp^2 and Si^2 represent the variances for each one of the split-halves of the test, and Sx^2, is the total variance of the test. The mean confidence level for all the students provides us with the reliability score of the test.

The reliability of the survey questionnaires was determined by using the Cronbach [13] coefficient α, in the expression:

$$\alpha = \frac{N}{N-1}\left[1 - \frac{\sum Si^2}{Sx^2}\right] \qquad (4)$$

Where α is the estimator of the coefficient of reliability, N is the number of test items, Si^2 is the variance of the subjects in relation to question i, and Sx^2 is the variance for the entire test.

The validity of the post-test is confirmed by comparing its results with those of other tests taken on the university course, thereby obtaining the correlations matrix shown in table 4, in which it can be seen that the correlations of the post-test are all over 0.2, except for the pre-test. This is entirely logical because, due to the students having no prior knowledge of pneumatics, the majority of them scored zero in the pre-test.

Correlations

		TEO-2º	MEAN_PO	PAR-2º	PAR-1º	LAM	JUNE	PRE
TEO-2º	Pearson's Correlation							
	Sig. (bilateral)							
	N							
MEAN_PO	Pearson's Correlation	.373**						
	Sig. (bilateral)	.000						
	N	138						
PAR-2º	Pearson's Correlation	.788**	.306**					
	Sig. (bilateral)	.000	.000					
	N	138	138					
PAR-1º	Pearson's Correlation	.528**	.357**	.560**				
	Sig. (bilateral)	.000	.000	.000				
	N	138	138	138				
LAM	Pearson's Correlation	.371**	.260**	.642**	.579**			
	Sig. (bilateral)	.000	.002	.000	.000			
	N	138	138	138	138			
JUNE	Pearson's Correlation	.485**	.430**	.552**	.535**	.322**		
	Sig. (bilateral)	.000	.000	.000	.000	.000		
	N	138	138	138	138	138		
PRE	Pearson's Correlation	-.062	.164	.099	-.011	.161	.095	
	Sig. (bilateral)	.467	.055	.249	.899	.059	.268	
	N	138	138	138	138	138	138	

**. The correlation is significant at a level of 0.01 (bilateral).

Table 4. Correlation matrix for scores in different tests

The lecture hall presentations, for groups T and A took place in the same lecture hall, with the same teacher and had a similar timetable. The theoretical concepts of pneumatics were explained in a one-hour session. A further one-hour session was used to identify elements and elaborate on simple pneumatic circuits and then came the final session, this time of two hours, in which the student was set the task of drawing and designing a pneumatic device. There was a period of one month for self-study, from the last lecture up until the post-test.

5. RESEARCH RESULTS

5.1. HYPOTHESIS 1. The pneumatic engineering design hypermedia as a medium through which to convey information is of an acceptable quality.

The first step was to translate the concepts referred to in hypothesis 1, into dimensions (EHi), from which the indicators were deduced (EHi.j), which then allowed us to measure the concepts being studied. The dimensions and indicators that were used are as follows:

EH1.- Aspect or appearance of screens. EH1.1 = colours and backgrounds used. EH1.2 = typesize and font. EH1.3 = design and positioning of navigational buttons. EH1.4 = quality of drawings. EH1.5 = quality of animations. EH1.6 = quality of audio.
EH2.- Arrangement of screen display. EH2.1 = screen display divided into lesson, sections, graphics area, and text area.
EH3.- User-friendly usage and navigation. EH3.1 = easy orientation. EH3.2 = easy navigation. EH3.3 = simple usage.
EH4.- Usefulness of the application in teaching. EH4.1 = to the lecturer in the lecture theatre. EH4.2 = as an additional tool in self-study.
EH5.- Usefulness of glossary, index and help. EH5.1 = Usefulness of the glossary, to find definitions of technical vocabulary. EH5.2 = usefulness of the index, to quickly find any concept on the hypermedia. EH5.3 = usefulness of the help function, to know how the hypermedia application works.

The Likert scale was used to measure the concepts, with the following categories and scores: very bad (1), bad (2), regular (3), good (4) and very good (5).

Univariable statistical analysis was used [1,14] to obtain the average $\bar{x}$, typical deviation S_{n-1}, the coefficient of asymmetry γ_1, the coefficient of kurtosis γ_2 and the typical mean estimation error $S_{n-1}/ n^{1/2}$, giving rise to the results shown in the table 5.

Dimensions	EH1.j						**EH1**	**EH2**	EH3.j	
Indicators	EH1.1	EH1.2	EH1.3	EH1.4	EH1.5	EH1.6		EH2.1	EH3.1	EH3.2
N valid	99	98	99	99	99	99	99	98	99	99
$\bar{x}$	3.9192	3.9592	4.0303	3.9798	3.9596	2.9394	3.7973	3.9286	3.9697	4.1010
S_{n-1}	.4842	.5546	.6617	.6368	.8797	.7398	.3569	.6462	.7065	.5975
$S_{n-1}/\sqrt{n}$	4.90E-02	5.60E-02	6.65E-02	6.42E-02	8.84E-02	7.44E-02	3.59E-02	6.53E-02	7.10E-02	6.00E-02
γ_1	-.742	-.391	-.248	-.223	-.471	.097	-.077	.067	-.312	-.327
γ_2	3.096	1.619	.052	.263	-.517	-.279	0518	-.563	.018	1.038
Dimensions	EH3.j	**EH3**	EH4.j		**EH4**	EH5.j			**EH5**	**ENC1**
Indicators	EH3.3		EH4.1	EH4.2		EH5.1	EH5.2	EH5.3		
N valid	99	99	99	98	99	95	96	94	97	**99**
$\bar{x}$	4.3838	4.1515	4.1515	4.2755	4.2121	3.7579	3.8333	3.6809	3.7629	**3.9717**
S_{n-1}	.6011	.4745	.8374	.8095	.6817	.6797	.6754	.6258	.5068	**.3379**
$S_{n-1}/\sqrt{n}$	6.04E-02	4.77E-02	8.42E-02	8.18E-02	6.85E-02	6.97E-02	6.89E-02	6.46E-02	5.15E-02	**3.97E-02**
γ_1	-.402	.040	-1.252	-.785	-1.058	-.696	-.205	.087	-.395	**-.386**
γ_2	-.651	.291	2.647	-.308	1.668	.829	.070	-.316	.749	**.688**

Table 5. Results of the survey used in the evaluation of hypothesis 1

In table 5, it can be seen that the only indicator with a grade below 3 is EH1.6. Enquiring into the nature of the problems students were encountering with the audio led to the conclusion that students had not put their headphones on correctly, that some students had switched on the microphone incorporated into their headphones, thereby creating interference, and that the volume on the sound cards of some computers was badly adjusted. All of these questions led to the insertion of a warning in the help menu of the application, which informs users to use the headphones with the microphone switched off, when studying on the hypermedia application.

It was also observed that the majority of the indicators presented a negative coefficient of asymmetry γ_1, which indicates a greater presence of values above the average, in the distribution of the indicators. It is worth noting the low typical error value of the survey, as well as its positive coefficient of kurtosis, γ_2.

The reliability of the questionnaire, arrived at by using the expression in (5), was 0.79.

In order to verify that the average grade of survey 1 (ENC1) was above 3, it was decided to carry out the following test: the statement that 'the average grade is lower or equal to 3 points' was taken as the null hypothesis; it was set against the alternative hypothesis, which was 'the average grade is above 3 points'. This test was performed using Student's t-test, which calculated the critical value at $p = 0.000$, meaning the average grade of ENC 1 was above 3. In order to check the degree to which the null hypothesis was false, the effect size was determined at a value of 2.875, which tells us that an average grade of more than 3 points was awarded by 99.8 % of the students in ENC1.

5.2. HYPOTHESIS 2. In relation to the proposed didactic objectives, the contents developed in the pneumatic engineering design hypermedia are acceptable.

The dimensions (ECi) and indicators (ECi.j) used in survey 2 (ENC2) are the following:

EC1.- Extension of the contents. EC1.1 = alignment of contents with objectives of the subject. EC1.2 = explanation of the concepts.
EC2.- Quality of the application. EC2.1 = depth of explanations. EC2.2 = rigour of explanations. EC2.3 = sequencing of the presentation of the contents. EC2.4 = alignment of resources (animations, audio, text, figures, etc.) with the explanations of the concepts.
EC3.- Clarity of explanations. EC3.1 = in the representation of pneumatic elements. EC3.2 = in the interpretation of pneumatic elements. EC3.3 = in the interpretation and working of the pneumatic elements.
EC4.- Usefulness of the application for technical drawing. EC4.1 = to represent and solve exercises step by step. EC4.2 = to simulate circuits using the Pneusim simulator. EC4.3 = in relation to the contents of the study unit: "pneumatic design and drawing". EC4.4 = in relation to the study unit on technical drawing.

The Likert scale has been used to measure the concepts, using the same categories and scores as in ENC1. The results of the univariable statistical analysis are shown in table 6, in which the symbols retain the same meaning as in table 5.

Dimensions	EC1.j		**EC1**	EC2.j				**EC2**	EC3.j
Indicators	EC1.1	EC1.2		EC2.1	EC2.2	EC2.3	EC2.4		EC3.1
N valid	99	98	99	99	99	99	99	99	99
$\bar{x}$	3.7172	3.7449	3.7374	3.4141	3.7475	4.0202	4.2424	3.8561	3.7172
S_{n-1}	.6707	.6629	.5409	.6702	.5596	.6848	.6403	.4257	.8576
$S_{n-1}/\sqrt{n}$	6.74E-02	6.97E-02	5.44E-02	6.74E-02	5.62E-02	6.88E-02	6.44E-02	4.28E-02	8.68E-02
γ_1	-.220	.117	.091	.113	-.366	-.415	-.263	-.146	-1.001
γ_2	.070	-.411	-.307	-.118	.295	.404	-.647	.499	1.723
Dimensions	EC3.j		**EC3**	EC4.j				**EC4**	**ENC2**
Indicators	EC3.2	EC3.3		EC4.1	EC4.2	EC4.3	EC4.4		
N valid	99	99	99	98	93	98	97	98	99
$\bar{x}$	3.6263	3.7980	3.7138	3.9286	3.9570	4.0918	4.1649	4.0357	3.8352
S_{n-1}	.8401	.9033	.6868	.8027	.8712	.6436	.6874	.5538	.411
$S_{n-1}/\sqrt{n}$	8.44E-02	9.08E-02	6.09E-02	8.11E-02	9.03E-02	6.50E-02	6.98E-02	5.94E-02	4.13E-02
γ_1	-.251	-.265	-.480	-.846	-.621	-.322	-.422	-.446	-.124
γ_2	.080	-.729	1.064	1.375	.341	.355	-.065	.009	.915

Table 6. Survey results used in the evaluation of hypothesis 2

From table 6, it may be seen that there are no indicators with grades of less than 3, that the majority of the indicators show a negative coefficient of asymmetry, and that the coefficient of kurtosis is positive.

The reliability of ENC2 has been calculated using the expression in (5), which gave a result of 0.82.

In order to verify that the average ENC2 grade was greater than 3, the following test was proposed:

- Null Hypothesis H_0: the average grade given by the students is below or equal to 3 points in ENC2.
- Alternative hypothesis H_1: the average grade given by the students is above 3 points in ENC2l.

This test was carried out using Student's t-test, giving a critical value of p = 0.00, which meant the proposed null hypothesis was rejected. In order to verify the degree to which the null hypothesis was false, the effect size was calculated at a value of 2.03, which tells us the ENC2 survey was, on average, graded with more than 3 points by 97.88% of the students.

5.3. HYPOTHESIS 3: Significant differences are found between the scores obtained by students prior to studying and after having studied the subject matter.

This test was only taken by the group with prior knowledge of pneumatics, since the majority of the other students scored zero in the pre-test, and it would not therefore have made sense to look for a relationship between the results of the pre-test and the post-test. To do so, students with prior knowledge were considered to be those who obtained a mark above or equal to 2.5 points in the pre-test. The composition of this group and its results are shown in table 7.

	N	Mean	Typical Deviation	Minimum Mark	Maximum Mark
pre-test	31	5.1703	1.8890	2.50	9.17
post-test	31	6.7540	2.1500	1.13	10.00

Table 7. Results of the group with prior knowledge

In the first place, an attempt was made to prove that the variation in the marks of the group in question, before and after having studied the material on pneumatic design and drawing, is positive, to a significant degree. To do so, the following hypothesis tests were proposed.

- H_0: Post-test score – Pre-test score = 0
- H_1: Post-test score – Pre-test score > 0

As they were the same students who took the pre-test and the post-test, the samples are related, for which reason Wilcoxon's [15] non-parametric test was used to determine the values. This test gave us the results shown in tables 8 and 9. As the critical value in Table 9 is p = 0.003, we had to reject the proposed hypothesis, which stated that the scores obtained by students with prior knowledge were equal before and after having studied the subject matter. In table 8, it may be seen that 22 students improved on their results in the pre-test, against 9 students who obtained a lower mark in the post-test than in the pre-test.

Ranges

		N	Mean range	Sum of all ranges
MEAN_PO-PRE	Negative Ranges	9 [a]	11.00	99.00
	Positive Ranges	22 [b]	18.05	397.00
	Ties	0 [c]		
	Total	31		

a. MEAN_PO < PRE

b. MEAN_PO > PRE

c. PRE = MEAN_PO

Table 8. The Wilcoxon signed rank test for the group of students with prior knowledge.

Test [b]

	MEAN PO - PRE
Z	-2.920 [a]
Asymptotic Sig(bilateral)	.003

a. Based on the negative ranges.

b. Wilcoxon signed rank test

Table 9. Significance level for the group of students with prior knowledge.

The relation between the results of the pre-test and the post-test, were measured using Spearman's coefficient of correlation r_s [15], obtaining a value of $r_s = 0.95$, which tells us that there is a strong positive relationship between the results of both tests.
To find out whether the size of r_s indicates a true relationship between the results of the pre-test and the post-test, the null hypothesis that the correlation coefficient of the population $\rho_{xy} = 0$ is tested. Using the t-statistic $T = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}}$ [15], which follows Student's t distribution to n –2 degrees of liberty, a critical value of p = 0.00 was chosen, on the basis of which the proposed hypothesis has to be rejected, or to put it another way, there is sufficient empirical evidence to confirm that there is a significant relation.

5.4. HYPOTHESIS 4: Those students who attend lectures and who have access to the hypermedia application perform better than those who only attend the lectures and those who only use the hypermedia.

The results of the univariable statistical analysis of the scores obtained by the students in the post-test is set out in table 10, from which the following information may be taken:

Descriptive

	SUBGROUP			Statistica	Typ. Error
MEAN_PO	A	Mean		6.5347	.280
		Confidence interval for a mean of 95%	Lower limit	5.9714	
			Upper limit	7.0981	
		for a mean reduced to 5%		6.5706	
		Median		6.6750	
		Variance		3.764	
		Typ. Dev.		1.9401	
		Minimum		1.12	
		Maximum		10.00	
		Range		8.88	
		Interquartile range		2.4292	
		Asymmetry		-.275	.34
		Kurtosis		.210	.67
	M	Mean		6.1723	.315
		Confidence interval for a mean of 95%	Lower limit	5.5378	
			Upper limit	6.8069	
		for a mean reduced to 5%		6.2089	
		Median		6.3917	
		Variance		4.671	
		Typ. Dev.		2.1612	
		Minimum		1.13	
		Maximum		10.00	
		Range		8.87	
		Interquartile range		3.0500	
		Asymmetry		-.162	.34
		Kurtosis		-.664	.68
	T	Mean		5.9516	.333
		Confidence interval for a mean of 95%	Lower limit	5.2795	
			Upper limit	6.6236	
		for a mean reduced to 5%		6.0693	
		Median		6.1167	
		Variance		4.769	
		Typ. Dev.		2.1838	
		Minimum		.30	
		Maximum		8.89	
		Range		8.59	
		Interquartile range		3.3250	
		Asymmetry		-.765	.36
		Kurtosis		.027	.70

Table 10. Univariable statistical analysis of the post-test scores in the different groups

- The average score of group (A), is 6% higher than the multimedia group (M), and 10% higher than the traditional group (T).
- The lower and upper limits of the mean, for a confidence interval of 95%, share a large common zone in the three groups, which ranges between 5.97 and 6.62 points.
- The variances in the three groups are similar and the coefficient of asymmetry is negative in all three cases.

The first step, in order to select the statistical test that would ascertain whether there are significant differences between the means of the three groups, was to check whether the distribution of the dependent variable (MEAN_PO) in the population followed a normal distribution, using the Kolmogorov – Smirnov test. This gave the results shown in table 11, in which the critical value is p = 0.099, on account of which the hypothesis of normality cannot be used, since there is only a probability of 9.9% that the score in the post-test has a normal distribution. As a result, any analysis of hypothesis 4 has to be performed using non-parametric tests.

Kolmogorov-Smirnov sample test

		MEAN_PO
N		138
Normal parameters [a,b]	Mean	6.2296
	Typ. Deviation	2.0928
Most extreme values	Absolute	.104
	Positive	.044
	Negative	-.104
Kolmogorov-Smirnov Z		1.227
Asymptotic sig. (bilateral)		.099

a. The test distribution is Normal.

b. Calculated on the basis of the data.

Table 11. Post-test normality table

In order to test whether there are significant differences in the mean scores of the three groups, the following hypotheses are postulated:

a) H_0: post-test mean for group A = post-test mean for group M
   H_1: post-test mean for group A ≠ post-test mean for group M
b) H_0: post-test mean for group A = post-test mean for group T
   H_1: post-test mean for group A ≠ post-test mean for group T
c) H_0: post-test mean for group A = post-test mean for group M = post-test mean for group T
   H_1: post-test mean for group A ≠ post-test mean for group M ≠ post-test mean for group T

Bearing in mind that the groups A, M and T are independent, and that the post-test score does not have a normal distribution pattern, the Mann – Whitney test is used to confirm a) and b), and the Kruskal – Wallis test to confirm c), which obtained the results shown in tables 12, 13 and 14.

Range **s**

SUBGROUP		Nos.	Average range	Sum of ranges
MEAN_PO	A	48	50.30	2414.50
	M	47	45.65	2145.50
	Total	95		

Statistical Tests **a**

	MEAN_PO
Mann-Whitney U Test	1017.5
Wilcoxon(Matched Pairs)	2145.5
Z-test	-.825
Asymptotic significance bilateral	.410

a. Grouping Variable: SUBGROUP

Table 12. Mann – Whitney test to compare groups A and M

Range **s**

SUBGROUP		Nos.	Average range	Sum of ranges
MEAN_PO	A	48	48.56	2331.0
	T	43	43.14	1855.0
	Total	91		

Statistical Tests **a**

	MEAN_PO
Mann-Whitney U Test	909.0
Wilcoxon(Matched Pairs)	1855.0
Z-test	-.981
Asymptotic significance bilateral	.326

a. Grouping Variable: SUBGROUP

Table 13. Mann – Whitney test to compare groups A and T

Range **s**

SUBGROUP		Nos	Average range
MEAN_PO	A	48	74.36
	M	47	67.91
	T	43	65.80
	Total	138	

Statistical Tests **a,b**

	MEAN-PO
Chi-squared	1.159
gl	2
Asymptotic significance	.560

a. Kruskal-Wallis test.

b. Grouping Variable: SUBGROUP

Table 14. Kruskal – Wallis test to compare groups A, M and T

After observing the critical values of the three tests, the conclusion is that the null hypothesis of equality of the mean post-test score for the three groups can neither be accepted nor rejected.

5.4.1. Analysis by objectives

An alternative to the statistical tests described above is to analyse the results of the three groups (A,M,T) in an independent way to compare the three didactic objectives of pneumatic design and drawing, which are described as follows:

1. Define the concepts of representation and symbolic interpretation of the pneumatic elements. This objective is measured by the variable “concept”.
2. Interpret the way the most important pneumatic elements function. The variable “function”.
3. Interpret and design pneumatic schematic diagrams. This objective is measured by the post-test variable “schematic diagram”.

By applying the Mann – Whitney test, the following information shown in tables 15 and 16 is obtained:

Ranges

SUBGROUP		Nos	Average range	Sum of ranges
CONCEPT	A	48	46.68	2240.50
	M	47	49.35	2319.50
	Total	95		
FUNCTION	A	48	54.75	2628.00
	M	47	41.11	1932.00
	Total	95		
SCHEMATIC DIAGRAM	A	48	47.58	2284.00
	M	47	48.43	2276.00
	Total	95		

Statistical Tests a

	CONCEPT	FUNCTION	SCHEMATIC DIAGRAM
Mann-Whitney U Tests	1064.500	804.000	1108.000
Wilcoxon (Matched Pairs)	2240.500	1932.000	2284.000
Z- test	-.527	-2.449	-.152
Asymptotic significance (bilateral)	.598	.014	.879

a. grouping variable: SUBGROUP

Table 15. Mann – Whitney test for groups A and M, by objectives.

Ranges

SUBGROUP		Nos	Average range	Sum of ranges
CONCEPT	A	48	46.97	2254.50
	T	43	44.92	1931.50
	Total	91		
FUNCTION	A	48	53.89	2586.50
	T	43	37.20	1599.50
	Total	91		
SCHEMATIC DIAGRAM	A	48	45.38	2187.00
	T	43	46.70	2008.00
	Total	91		

Statistical Tests a

	CONCEPT	FUNCTION	SCHEMATIC DIAGRAM
Mann-Whitney U Tests	985.500	653.500	1002.000
Wilcoxon (Matched Pairs)	1931.500	1599.500	2178.000
Z- test	-.411	-3.061	-.244
Asymptotic significance (bilateral)	.681	.002	.808

a. grouping variable: SUBGROUP

Table 16. Mann – Whitney test, by objectives, for groups A and T.

The critical value $p < 0.05$ of the variable “function”, between groups A – M and A – T, tells us to accept the alternative hypothesis, stating that the academic performance of group A differs from that of the other two groups.

The critical value of the variable “concept”, does not allow us to accept or reject the null hypothesis of equality for this variable.

The critical value $p>0.8$ of the variable “schematic diagram”, indicates that we must accept the hypothesis of equality of the means in relation to the academic performance of groups A, M and T.
The same conclusions may be reached from the Kruskal – Wallis test for the objectives of the three groups, as shown in table 17

Ranges

SUBGROUP		Nos	Adverage range
(SCHEMATIC) DIAGRAMS	A	48	68.46
	M	47	69.72
	T	43	70.42
	Total	138	
CONCEPT	A	48	69.15
	M	47	73.04
	T	43	66.02
	Total	138	
FUNCTION	A	48	84.14
	M	47	63.71
	T	43	59.49
	Total	138	

Statistical tests a,b

	SCHEMATIC DIAGRAMS	CONCEPT	FUNCTION
Chi-squared	.059	.864	10.455
gL	2	2	2
Asymptotic significance	.971	.649	.005

a. Kruskal-Wallis test
b. Grouping variable: SUBGROUP

Table 17. Kruskal – Wallis test by objectives on the three groups A, M and T

5.5. HYPOTHESIS 5: The influence of the hypermedia is greater in those students with less intellectual aptitude than in those students with greater intellectual aptitude.

As a first step, the concepts greater and less were defined. Students with less intellectual aptitude (LESS/APT) are those whose score f(SR,MR) was below the mean average of the sample. The remainder were students who had a greater intellectual aptitude (GREATER/APT).

In the same way, the average score of the post-test sample divided the students into two groups. Those with HIGHER/PERFORM, who scored above the average, and those with LOWER/PERFORM, who scored below the average.

Two procedures were used to study the relationship between intellectual aptitudes and academic performance:

- Evaluating the number of students belonging to each category
- Evaluating the increase in academic performance

5.5.1. Evaluating the number of students belonging to each category

For this study, the results from groups A and M were used, as what we are interested in studying is whether the hypermedia exercises greater influence on students with less intellectual aptitude.

Table 18 shows the absolute frequencies obtained for each combination of the values of the variables.

	MEAN POST-TEST		$n_{i.}$
	HIGHER/PERFORM	LOWER/PERFORM	
GREATER/APT	31	20	51
LESS/APT	20	24	44
$n_{.j}$	51	44	95

Table 18. Contingency MEAN_PO table and intellectual aptitudes for groups A and M

The null hypothesis, stating the independence of academic performance and intellectual aptitudes is tested against the alternative hypothesis, stating that both variables are linked.

The statistic X^2 takes a value of 2.23, which in the Pearson distribution χ_1^2 assumes a critical value of p = 0.15. This does not allow us to either accept or reject the hypothesis of independence between the academic performance and the intellectual aptitudes of the students.

5.5.2. Evaluating the increase in academic performance

The variable that defines an increase in academic performance was taken as the difference between the score obtained in the post-test and the pre-test.

INCREASE = MEAN_PO – MEAN PRE-TEST

The mean results of these variables, in table 19, which were analysed using the Mann – Whitney test, gave us the results as shown in tables 20 and 23.

	All	All		Support subgroup		Multimedia subgroup		Traditional subgroup	
		Less/app	Greater/apt	Less/app	Greater/apt	Less/app	Greater/apt	Less/app	Greater/apt
Number	137	73	64	24	23	26	21	23	20
Mean	4.89	4.5	5.34	5.18	5.58	4.75	4.75	3.51	5.67
Typ. Dev.	2.7	2.52	2.58	1.93	2.87	2.6	3.2	2.72	2.39
Typ. Mean Err.	0.23	0.29	0.35	0.38	0.59	0.51	0.71	0.56	0.53

Table 19. Average increase in academic performance.

Ranges

	f(SR,MR)	Nos	Adverage range	Sum of ranges
INCREASE	LESS/AP	73	63,14	4609,00
	GREATER/AP	64	75,69	4844,00
	Total	137		

Statistical tests [a]

	INCREASE
Mann-Whitney U Test	1908,000
Wilcoxon (Matched Pairs)	4609,000
Z	-1,847
Asymptotic significance (bilateral)	,065

a. grouping variable: f(SR,MR)

Table 20. Man-Whitney test for all students.

Ranges

	f(SR,MR)	Nos	Adverage range	Sum of ranges
INCREASE	LESS/AP	24	22,52	540,50
	GREATER/AP	23	25,54	587,50
	Total	47		

Statistical tests [a]

	INCREASE
Mann-Whitney U Test	240,500
Wilcoxon (Matched Pairs)	540,500
Z	-,756
Asymptotic significance (bilateral)	,450

a. grouping variable: f(SR,MR)

Table 21. Man-Whitney test for students in the support subgroup.

Ranges

	f(SR,MR)	Nos	Adverage range	Sum of ranges
INCREASE	LESS/AP	26	24,04	625,00
	GREATER/AP	21	23,95	503,00
	Total	47		

Statistical tests [a]

	INCREASE
Mann-Whitney U Test	272,000
Wilcoxon (Matched Pairs)	503,000
Z	-.021
Asymptotic significance (bilateral)	,983

a. grouping variable: f(SR,MR)

Table 22. Man-Whitney test for students in the multimedia subgroup.

Ranges

	f(SR,MR)	Nos	Adverage range	Sum of ranges
INCREASE	LESS/AP	23	17,63	405,50
	GREATER/AP	20	27,02	540,00
	Total	43		

Statistical tests [a]

	INCREASE
Mann-Whitney U Test	129,500
Wilcoxon (Matched Pairs)	405,500
Z	-2,447
Asymptotic significance (bilateral)	,014

a. grouping variable: f(SR,MR)

Table 23. Man-Whitney test for students in the traditional subgroup.

In the overall study for all students, table 20, a critical value of p=0.065 was obtained, which tells us that the students with greater intellectual aptitude are those who obtain a greater increase in academic performance.

The critical value of the support group, table 21, is p = 0.450, which does not lend weight to equality or inequality of an increase in academic performance between students with greater or less intellectual aptitude, despite the average increase in the group with greater intellectual aptitude being 5.98 points, as against 5.18 points for students with less aptitude.

In the multimedia group, in table 22, a critical value of p = 0.983 is obtained, which allows us to affirm that in this group there is equality between groups with greater or lesser intellectual aptitude as regards the increase in academic performance.

The critical value for the traditional group, in table 23, is p = 0.014, which allows us to state that the increase in academic performance is greater among students with greater intellectual aptitude.

5.6. HYPOTHESIS 6: Regardless of any prior knowledge and the capacity of the student, in the self-study group, the academic performance is unrelated to the study time spent on the hypermedia application

This hypothesis is proposed solely for the multimedia group (self-study) group, given that in the support group, students might occasionally use the hypermedia application to study certain sections without studying an entire set of lessons, which might distort the relationship between length of study time on the application and academic performance.

As for hypothesis 5, the variables are divided up into two categories. For the time-related variable, those students with the least study time are those who spent less time studying than the mean average of the sample, the remainder being those students whose study time was above the average. Likewise, the average score of the sample in the post-test, divides the students into two categories of better and worse academic performance.

The null hypothesis, stating that the two variables, academic performance (MEAN_PO) and length of study time (TIME) are not associated was tested against the alternative hypothesis, stating that an association exists between them.

Table 24 shows the absolute frequencies obtained for each combination of the different categories of variable.

	MEAN_PO		$n_{i.}$
	BETTER/PERFORM	WORSE/PERFORM	
GREATER/T	15	4	19
LESS/T	10	19	29
$n_{.j}$	25	23	48

Table 24. Academic performance-length of study time contingency table

The Chi-squared test X^2 gives a critical value of p = 0.003, which leads us to reject the hypothesis of independence and accept that there is an association between the length of study time and academic performance. The direction and intensity of the association was obtained by using Yule's Q factor [14] that gave a value of 0.75, which tells us that there is a strong relationship between the length of study time and the academic performance of the student.

6. CONCLUSIONS

In the introductory section, it was stated that this research was intended to reveal the quality of the “pneumatic engineering design drawing” hypermedia, as well as the effectiveness and efficiency of the teaching-learning process of pneumatics. The time has now come to put forward the most important findings of the research.

The quality of the application was measured by two surveys, one to evaluate its quality as a means of conveying information, and another to evaluate the degree to which its contents are aligned with the proposed didactic objectives. As a means of conveying information, it scored 3.97 points out of 5, with a typical error of 0.034, meaning that 99.8% of students awarded the application a grade of more than 3 points. With regard to the alignment of its contents with the objectives, it scored 3.83 points, with a typical error of 0.041, meaning that 97.88% of students graded this section with a mark of more than 3 points. According to the scale used, it may be said that the quality of the application is good, a grading that is not markedly different from other research results [4,5].

The effectiveness of the application was measured in terms of increased knowledge among students with prior knowledge of the subject matter, the academic performance of the different groups, and the influence of the hypermedia, for the different types of students, on the learning process.

In the group of students with prior knowledge, the results of the post-test are strongly related to the results of the pre-test, undergoing an average increase of 1.6 points in the test results.

Despite the average score of the support group being 6% higher than the multimedia group and 10% higher than the traditional group, there is insufficient empirical evidence to affirm that the students who attend lectures and who also rely on the hypermedia perform better than those who only attend lectures, or those who study the subject by themselves with the help of the hypermedia application. It is only in the section on interpreting the functions of pneumatic elements that the support group shows any significant difference with respect to the multimedia group and the traditional group.

This result is of great importance, since the implementation of the ECTS credits will make it necessary to identify those parts of the subject matter that the student can study privately with the help of suitable material and with the assistance of tutorials. The results demonstrate that this part meets the aforementioned condition.

With regard to the relationship between academic performance and intellectual aptitudes among the students who used the hypermedia, the study based on the number of students in each category does not allow us to establish a link between academic performance and intellectual aptitudes although a tendency towards a positive association is observable, which could only be further justified by carrying out more tests. However, bearing in mind the increase in the post-test scores compared with the pre-test scores in the multimedia group, the independence of academic performance with respect to intellectual aptitude can be affirmed; whereas in the traditional group, the students with the greatest intellectual aptitude are those who perform best.

The efficiency of the application was measured in terms of the relationship between the length of study time and academic performance, and by monitoring the length of time the students in the multimedia group spent on the application.

There is a strong positive relationship between the time spent on the application and the academic performance of the student.

In relation to the variable for study time on the application, an average of 93 minutes was recorded, with a typical deviation of 45 minutes, which when set against the 4 hours used to develop the materials in the support group and traditional group, implies a minimum time saving of 42%. This figure does not take into account study time by students from these two groups spent studying the subject matter privately. This is in accordance with results obtained in other studies [3,4].

BIBLIOGRAPHY

[1] Cea D'Ancona Mª.A. *Metodología cuantitativa .Estrategias y Técnicas de investigación* .Madrid. Síntesis 1996

[2] McDonald S. David. "The influence of multimedia training on users' attitudes: lessons learned". Computer & Education, 42. 2004.

[3] Schmeeckle M. Joyce. "Online training: An evaluation of the effectiveness and efficiency of training law eforcement personnel over the internet". Journal of Science Education and Technology". Vol 12. No 3. 2003.

[4] Peláez Vara, J. Álvarez de Ron González, V. Pascua Frades, J. "Evaluation of the multimedia machine tools and mechanized processes". International Journal of Computer Applications in Technology". Vol 14. Nos 1/2/3. 2001.

[5] Wu-Yuin Hwang. Nian-Shing Chen. Rueng-Lueng Hsu. "Development and evaluation of multimedia whiteboard system for improving mathematical problem solving". Computer & Education 2004.05.005

[6] Joffrey R. Appling and Lisa C. Peake. "Instructional technology and molecular visualization". Journal of Science Education and Technology. Vol 13. No 3. September 2004.

[7] Regina Schoenfeld-Tacher, Loretta L. Jones, and Kay A. Persichitte. "Differential effects of a multimedia goal-based scenario to teach introductory biochemistry – Who benefits most?". Journal of Science Education and Technology. Vol 10. No 4. December 2001.
[7] Regina Schoenfeld-Tacher, Loretta L. Jones, and Kay A. Persichitte. "Differential effects of a multimedia goal-based scenario to teach introductory biochemistry – Who benefits most?". Journal of Science Education and Technology. Vol 10. No 4. December 2001.

[8] Eric Tucker. "The impact of CAL in a university engineering school". Education + Training. Vol 12. No 7. 1997

[9] Bennett, G.K. et al. *DAT-SR. Forma L (cuadernillo de test).* Madrid. TEA ediciones. 1971.

[10] Bennett, G.K. et al. *DAT-MR. Forma L (cuadernillo de test).* Madrid. TEA ediciones 1972.

[11] Failes Baz, R. *Eficacia de un sistema AEO hypermedia en la enseñanza de dibujo técnico industrial.* Tesis doctoral. Bilbao. Universidad del País Vasco, ETSII y de ingenieros de telecomunicación, 1999.

[12] Matas Terrón A. "Differential análisis of indicators of achievement in objective tests". Revista electrónica de investigación y evaluación educativa. Vol 9. No 2. 2003.

[13] Muñiz, J. *Teoría clásica de los tests.* Madrid. Pirámide. 1998

[14] Ruiz Maya, L. Martín Pliego, J.F. Montero, J.Mª. Uriz Tomé, P. *Análisis estadístico de las encuestas: datos cualitativos.* Madrid. AC. 1995.

[15] Pardo A. y San Martín R. *Análisis de datos en Psicología II.* Madrid. Pirámide. 1998.