

**Francisco J. Almeida-Martínez
Jaime Urquiza-Fuentes**

**Long term evaluation using
VAST.**

Number 2011-13

**Serie de Informes Técnicos DLSI1-URJC
ISSN 1988-8074
Departamento de Lenguajes y Sistemas Informáticos I
Universidad Rey Juan Carlos**

Index

1. Introduction	3
2. Description of the evaluation	3
2.1. Experimental design, subjects, tasks and protocol	3
3. Results	4
3.1. Analysis of the marks obtained in the practices	4
3.2. Implication of students	5
4. Conclusions	5
References	6

Long term evaluation using VAST

Francisco J. Almeida-Martínez and Jaime Urquiza-Fuentes

Universidad Rey Juan Carlos
Departamento de Lenguajes y Sistemas Informáticos I
c/ Tulipn s/n, 28933 Madrid, Spain
francisco.almeida@urjc.es, jaime.urquiza@urjc.es

Abstract. This evaluation was carried out using the results of the last three evaluations. As the students in previous evaluations participated in the same group, they were used to work with VAST. We used an obligatory partial delivery of the practice of the subject to analyze in long term the effects of using VAST. The results of this study are satisfactory. We observed that the number of deliveries is higher in students who used VAST and also their marks are higher.

1 Introduction

The results of previous evaluations [1] have been satisfactory from the educational effectiveness and usability-quality points of view. From these results we observed positive results after using VAST during a long period of time. Due to this we decided to performed another long term evaluation. In this case this evaluation was formed by three short term evaluations [2], [3] and [4].

The objective of this evaluation is to see if after using VAST during a long period of time, students are more motivated in the subject, improving their motivation and definitely their learning.

2 Description of the evaluation

In this section we describe the evaluation. We refer to the participants, the experiment's design, the tasks performed during the session and the protocol.

2.1 Subjects

In this study we took into account **40 students** (all who delivered the practice) of the subject *Language processing* at the Rey Juan Carlos University during the 2010-2011 course.

2.2 Experimental design, tasks and protocol

The objective of this analysis is to study the students' implication in the subject after using or not VAST during a long period of time. Due to this we performed

three evaluations [2], [3] and [4] which allowed to obtain independent results about the short term evaluation of the tool.

Once the students participated in different evaluations, they had to deliver (optionally) the syntax analysis of the final practice. This delivery was based incentive in a 8% over the final mark in the practice. For this delivery we specified 16 criterials in order to correct the exercise: solution of ambiguities, operators' precedence, semantic actions, tests examples, etc. After correcting the exercises all students had just a numeric mark. We separated the students in three groups (G0-G2). The group *G0* was formed by students who participated in the control group in all the previous evaluations. *G1* was formed by students who participated in the treatment group in all the previous evaluations. Finally, *G2* was formed by students who did not participate in all the previous evaluations. Take into account that the practice could be done in group up to 3 studetns. In table 1 we show the protocol followed in this evaluation.

Control	Treatment
	LL short term evaluation
	LR short term evaluation
	Syntax error recovery short term evaluation

Table 1. Protocol followed in this evaluation

3 Results

In this section we describe the results of the evaluation. This evaluation had two main objectives. On the one hand we observed if the marks between both groups were significant different. On the other hand, we analyzed the implication of the students in the subject.

3.1 Analysis of the marks obtained in the practices

After building the different groups we performed a Kruskal-Wallis analysis which did not show any statistical significant difference ($X^2=1.13$, $p=0.77$). Afterwards we performed an analysis considering only pair of groups: G0-G1, G0-G1, G0-G2. This one neither show any difference between the groups. However, after studying the results of the analysis between G0-G1 ($U=22.50$, $p=0.07$), which average marks were 7.25 and 11.27 respectively, we can observe a big difference between marks.

Besides, this analysis showed that the minimum and maximum marks obtained in each group were different. The marks in G0 and G2 varied between 6-16 while in G1 it varied between 11 and 16. Finally we did a qualitative analysis segmenting the marks in four groups: D, C, B, A.

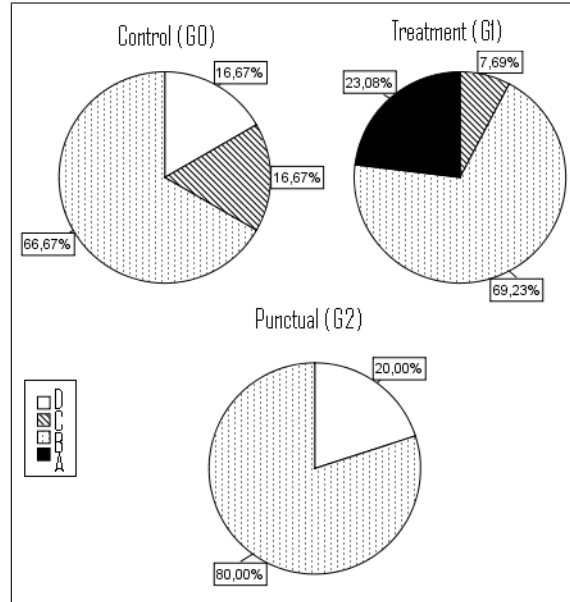


Fig. 1. Qualitative results

3.2 Implication of students

We studied the implication of students analyzing the number of deliveries according to the participation in previous evaluations. We considered 24 deliveries of this partial practice. We observed that the 54.27% (13/24) of the students who delivered the practice had used VAST in previous evaluations. The 25% (6/24) had used ANTLR/JFlex-Cup. Finally, 20.83% (5/24) had used just one tool.

If we focus on students who did not deliver the practice (16 students) we observed that 37.5% (6/16) had use just once VAST; 12.5% (2/16) had used ANTLR/JFlex-Cup. Only one students participated in the three previous evaluations using VAST and did not deliver the practice. The rest of students 43.75% (7/16) did not participated in any evaluation.

4 Conclusions

The results obtained in this long term evaluation allow to distinguish two fundamental aspects: the learning improvement and the students' implication.

Referring to the learning improvement, the marks obtained in the practices show that there does not exist any different between the groups. This means that VAST did not affect the mark in the partial delivery. However, after analyzing the marks between the control and treatment group (G0-G1), we observed that the average marks of the treatment group were much higher than the control one.

This could indicate that VAST may help students to implement the practice. We also observed that the minimum and maximum marks varied in a wide range being higher in treatment group (VAST). As it can be observed in Fig. 1 students who used VAST got, in general, better grades than the others. None of the students who used VAST failed the practice. This results show that although there are not significant differences between groups, VAST helps students to perform better their practices, so they get better marks.

According to the students' implication, we observe that the number of deliveries is higher in those students who used VAST in previous evaluations. Due to this we can say that students who used VAST were more motivated in the subject.

5 Acknowledgment

This project is supported by project TIN2008- 04103/TSI of the Spanish Ministry of Science and Innovation.

References

1. Francisco J. Almeida-Martínez and Jaime Urquiza-Fuente. Evaluación educativa de VAST con analizadores LR(1). Visualización vs no visualización. Technical Report 2010-04, Universidad Rey Juan Carlos, 2010.
2. Francisco J. Almeida-Martínez and Jaime Urquiza-Fuente. Building LR(1) grammars and parsers using VAST and Cup. Technical Report 2011-11, Universidad Rey Juan Carlos, 2011.
3. Francisco J. Almeida-Martínez, Jaime Urquiza-Fuente, and Antonio Pérez-Carrasco. Syntax error recovery visualization. evaluation using VAST and Cup. Technical Report 2011-11, Universidad Rey Juan Carlos, 2011.
4. Francisco J. Almeida-Martínez, Jaime Urquiza-Fuente, and Antonio Pérez-Carrasco. Building LL(1) grammars and parsers using VAST and ANTLR. Technical Report 2011-06, Universidad Rey Juan Carlos, 2011.