

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/311313792>

Characteristics of self-regulation of engineering students to predict and improve their academic performance

Conference Paper · October 2016

DOI: 10.1109/FIE.2016.7757479

CITATIONS

4

READS

80

4 authors:



Julieta Noguez

Tecnológico de Monterrey

77 PUBLICATIONS 467 CITATIONS

[SEE PROFILE](#)



Luis Neri

Tecnológico de Monterrey

49 PUBLICATIONS 280 CITATIONS

[SEE PROFILE](#)



Andrés González-Nucamendi

Tecnológico de Monterrey

11 PUBLICATIONS 19 CITATIONS

[SEE PROFILE](#)



Víctor Robledo-Rella

Tecnológico de Monterrey, Campus Ciudad de México

31 PUBLICATIONS 83 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Spatially Integrated Spectroscopy of Galactic HII Regios [View project](#)



Open Innovation Laboratory [View project](#)

Characteristics of self-regulation of engineering students to predict and improve their academic performance

Julieta Noguez^{1,2}, Luis Neri^{1,2}, Andres González- Nucamendi^{1,3}, Víctor Robledo-Rella^{1,3}

¹ Tecnológico de Monterrey, Campus Ciudad de Mexico

² Escuela de Educación, Humanidades y Ciencias Sociales

³ Escuela de Diseño, Ingeniería y Arquitectura
{jnoguez, neri, anucamen, vrobledo}@itesm.mx

Abstract—This paper describes the adaptation and validation of an instrument aimed to determine self-regulation skills, learning strategies and affective strategies of engineering students from the Tecnológico de Monterrey, Mexico City Campus. A statistical validation with Cronbach's alphas and a social validation in which students indicated whether or not they agree with their results on each dimension were carried out. An online system where students answered an appropriate questionnaire to determine their *student-profile* was developed. The *social validation* shows a very good agreement among the profiles obtained with the instrument and the perception of the students. Suitable statistical techniques allow classifying samples of students in different clusters according to their profiles, where members of each cluster have similar profiles. Finally, this instrument, along with additional student academic information, allows to predict their academic performance based on statistical methods, and provides support for instructors to focus their teaching strategies and methods, in order to work and pay attention on those students whose self-regulation dimensions were relatively poor.

Keywords—*self-regulation, learning strategies, affective strategies, statistical validation, social validation*

I. INTRODUCTION

Self-regulation refers to the ability of the student to take the most appropriate or optimal learning decisions. Along with emotional and relevant learning strategies, self-regulation may play an important role in academic performance. One factor that certainly impacts student learning is their ability to regulate themselves, that is, to recognize their own abilities and academic skills in order to stay focused on achieving a particular goal.

The Cyberlearning and Data Science Research Group at the Tecnológico de Monterrey, Mexico City Campus is carrying out a big-scope project aimed to study the main factors that could determine academic outcomes for students enrolled in engineering majors [1]. The study includes assessing *a)* student learning styles (LS), *b)* multiple intelligences (MI), and *c)* self-regulation skills (SR), as well as affective and learning strategies, in order to build a general *student profile*. Preliminary discussions of the instruments adopted to measure these constructs and partial results have been presented by [2], [3] and [4], respectively. This paper focuses on the adaptation and validation of the instrument aimed to determine self-regulation

skills (SR), learning strategies (LeSt) and affective strategies (AfSt) of engineering students at the Tecnológico de Monterrey, Mexico City Campus. A statistical and data mining efforts in order to understand the relation among the different dimensions of the instrument and the academic performance of the students were also carried out.

In order to study in detail these skills and strategies, the items proposed by Gargallo [5] were selected. It was determined that there should be a work of reorganization and adaptation of Gargallo's items in order to identify those relevant dimensions that could help instructors to characterize students' profiles. Therefore, the items were grouped and 8 dimensions were defined, namely: *i)* intrinsic motivation, *ii)* extrinsic motivation, *iii)* fitness and mood, *iv)* anxiety, *v)* self-regulation, *vi)* social interaction, *vii)* searching and selection of information strategies, and *viii)* processing and use of information strategies. A reliability study and a social validation of the dimensions of the adapted instrument were also performed. This adapted instrument is expected to be useful in order to construct an initial inventory of the above dimensions, so that both teachers and students could use it as an input to propose suitable strategies aimed to improve academic performance.

II. THEORETICAL FRAMEWORK

Several authors have discussed the elements that can have an impact to improve student learning, in order to guide teachers to rethink their teaching strategies to make them more effective. This is not an easy task because there are many and varied factors that may influence academic student performance, which range from student's intrinsic features, such as cognitive skills and personality, to external factors, such as the use of appropriate learning environments, adequate motivation strategies, or the adoption of suitable learning methodologies, among other.

One factor that certainly impacts student learning is their self-regulation skills, that is, the capability to know their own abilities and academic skills and to use them in order to stay focused to achieve a particular goal [6, 7]. In learning environments, it is desirable that students, in addition to performing learning activities, maintain a sense of *self-efficacy* to learn, assess their own learning, and keep the belief that they

will get positive results, maintaining a positive attitude and enjoying what they are doing.

There are different theories that suggest that the process for classifying and assessing progress in the categories of self-regulation of a student is a diagnostic process that can evolve according to their performance [7]. All these theories share a common ground where self-regulation is composed of different aspects (e.g. monitoring, goals setting, etc.). In addition to being cyclical, i.e., involving the interaction of personal, behavioral and environmental factors that change during the learning process, these components can be supervised to lead to desirable changes in strategies, cognitions, emotions and behaviors of the learner.

Following these theories, several instruments designed to assess self-regulation skills have been proposed. Examples are included in Fernandez for university students in Lima, Peru [8], Caso et al. for junior high school students in Baja California, México [9], and Gargallo et al. for university students in Valencia, Spain [5]. However, because these instruments were designed for the academic level and culture of the region where the students belong, these instruments cannot be directly applied to determine the skills and strategies of other target students, but must be adapted to their particular characteristics [10].

Due to the fact that Gargallo's questionnaire was also intended for college students, as it is the case for our student sample, and contains more items than other questionnaires (88), it was chosen and adapted to assess students SR, LeSt and AfSt. Therefore, the dimensions of the resulting instrument include self-regulation skills, learning strategies, affective and motivational elements, as well as cognitive and meta-cognitive elements. The instrument emphasizes the strategic use of driving components, such as awareness, intentionality, flexibility, supervision skills and self-regulation skills.

III. PROBLEM STATEMENT

The main objectives of this paper are the following:

- i) To validate statistically and socially the adapted instrument in order to properly assess the self-regulation skills (SR), learning strategies (LeSt) and affective strategies (AfSt) of our student sample.
- ii) To classify students in groups or clusters with similar student profiles, according to their SR, AfSt and LeSt dimensions.
- iii) To use the clusters of student profiles to study the relationship among student academic outcomes and their cognitive and social profiles. This would provide teachers with useful information in order to implement those educational activities that best suit the profiles of their students.

IV. METHODS AND RESULTS

We worked with a sample of $N = 96$ engineering students enrolled in Mathematics, Physics and Computing courses at the Tecnológico de Monterrey, Mexico City Campus.

The main stages of the method applied in this work are shown in Fig. 1, and they are described below.

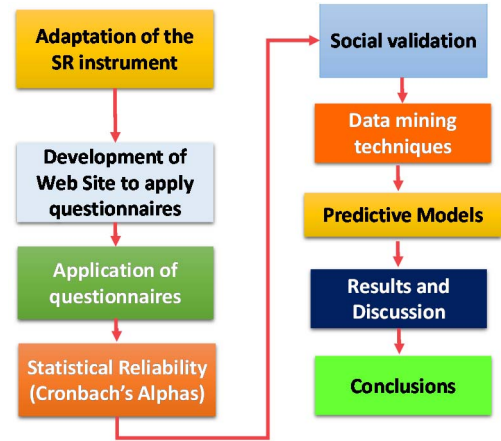


Figure 1. General method applied

A. Adaptation of the SR instrument

As we mentioned before, Gargallo's questionnaire [5] was adapted in order to assess student SR, LeSt and AfSt. However, it was necessary to review and change the wording of some items to adapt them to the culture of Mexican students. The 88 items conforming the questionnaire were regrouped in the 8 categories listed in Table I, where each category comprises a set of related items measuring the same dimension.

TABLE I. SELF-REGULATION DIMENSIONS

Short name	Name
<i>IntMot</i>	Intrinsic motivation
<i>ExtMot</i>	Extrinsic motivation
<i>Mood</i>	Fitness and mood
<i>Anx</i>	Anxiety
<i>SelfReg</i>	Self-regulation
<i>SocInt</i>	Social interaction
<i>InfSearch</i>	Strategies to search and select of information
<i>InfProc</i>	Strategies to use and process information

In order to obtain numerical values for each of the 8 dimensions in Table I, a Likert scale of 5 levels for each dimension was calculated, where 5 stands for strong agreement and 1 for strong disagreement. Thus, the strength of each dimension was calculated from the average of the set of items assigned to that dimension [11]. This results in a number between 1 and 5 for each dimension. If the value lies between 1 and 3 it is considered that the dimension is poorly developed, while if it lies between 3 and 5, the dimension is well developed. Finally, the profile for each student was obtained considering the values of the 8 dimensions.

B. Implementation of the instrument

Once the questionnaire was adapted, it was uploaded to an online system so that the students could answer it easily, and the

information could be processed automatically. The URL for the questionnaire online is:

<http://elearning2.ccm.itesm.mx/Encuestas/Alumnos/index.php>, (Fig. 2).

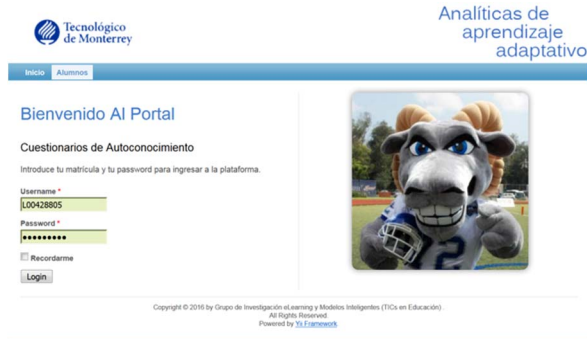


Figure 2. Our online Web system to apply self-regulation questionnaires

C. Application of questionnaires

During the August-December 2014 semester (AD2014) a sample of $N = 96$ engineering students, from 5 sections of Physics, Mathematics and Computer Science, was requested to answer the online questionnaire. After answering the questionnaire, each student got his/her profile, formed from the average values of each dimension. The results were given to the students both numerically and in a radar diagram, as shown in Fig. 3.

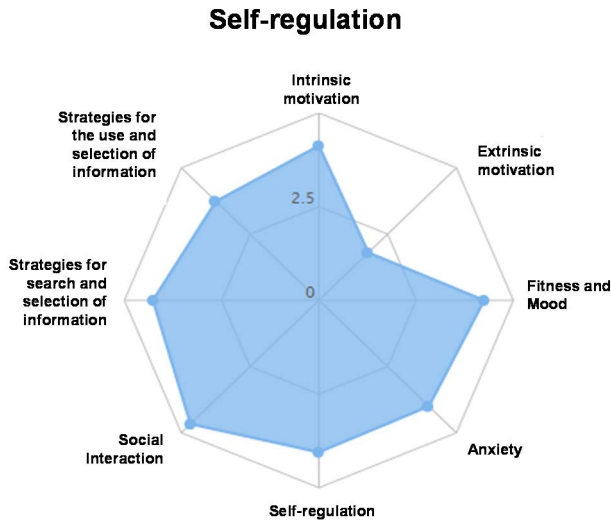


Figure 3. Example of a student profile showing the values for each dimension in a radar diagram. The scale ranges from 0 (center) to 5 (extremes).

After obtaining their profiles, students were asked to review the meaning of each dimension, to analyze the results given by the system, and to answer whether they agreed or not with the given score. We analyzed this comparison along with particular comments given by the students. An example of the explanation given to students for 3 of the 8 dimensions of Fig. 3 is shown in Table II. An example of the opinion given by a student on her results for those same dimensions is given in Table III.

TABLE II. EXAMPLE OF EXPLANATION OF THREE DIMENSIONS

Self-Regulation and Learning Strategies Dimension	Definition and Characteristics
Intrinsic Motivation	Degree to which the student is intrinsically motivated, committed and self confident to study with enthusiasm
Extrinsic Motivation	Degree to which the student depends on external factors in order to focus to study
Physical and emotional state	Degree to which the student maintains appropriate fitness and emotional state in order to obtain a better academic performance.

TABLE III. STUDENT'S PERCEPTION AND COMMENTS ON THE RESULTS OBTAINED FOR THE DIMENSIONS GIVEN IN TABLE II.

Result	Do you agree? YES or NO	Comments
4.125	Yes	In general, the motivation has to be promoted by myself
1.8	Yes	In general I study by myself and not for someone else.
4.25	Yes	I tend to be a happy person and seek to incorporate this to work, study, etc.

Once the results for our student sample for the 8 dimensions were obtained, a reliability study for the instrument was carried out, as shown in the next section.

D. Statistical Reliability (Cronbach's Alpha)

Given that all our students are from engineering majors in the first semesters of their careers, we decided to analyze the reliability of the instrument using the whole sample. Cronbach's alpha values [12] were therefore calculated for each dimension. The results are given in Table IV along with the number of items conforming each dimension. Six of the dimensions are internally consistent and two of your dimensions are inconsistent, i.e. six of the dimensions had values greater than 0.700, except for the dimensions Anx (0.593) and ExtMot (0.681), which are not too far from being statistical consistency, and have to be further investigated. Overall, we consider the adapted instrument is statistically reliable.

TABLE IV. CRONBACH'S ALPHA FOR EACH DIMENSIONS

Dimension	# items	Cronbach's alpha
IntMod	15	0.953
ExtMot	5	0.681
Fitness and Mood	4	0.816
Anx	4	0.593
SelfReg	19	0.914
SocInt	6	0.876
InfSearch	8	0.739
InfProc	27	0.924

E. Social validation

Subsequently, in order to evaluate the perception of students with respect to the values obtained in their profiles in each dimension, a *social validation* of the instrument was performed. As mentioned above, the students answered for each dimension whether they agreed or not with the value assigned by the system. A 95% concordance among the results given by the instrument and the perception of students was obtained.

F. Data mining techniques

From the information obtained above, an exploratory study with three statistical techniques was carried out: *a) Principal Components Analysis (PCA)* to obtain the diagram of principal components for the 8 dimensions, *b) Correlation Analysis (CA)* to obtain the correlation matrix for the 8 dimensions, and *c) Clustering* to group student profiles with similar characteristics. The plots derived with the R statistics package are shown next.

1. Principal Components Analysis (PCA).

The Principal Components Analysis diagram for the 8 dimensions considered in this work is shown in Fig. 4.

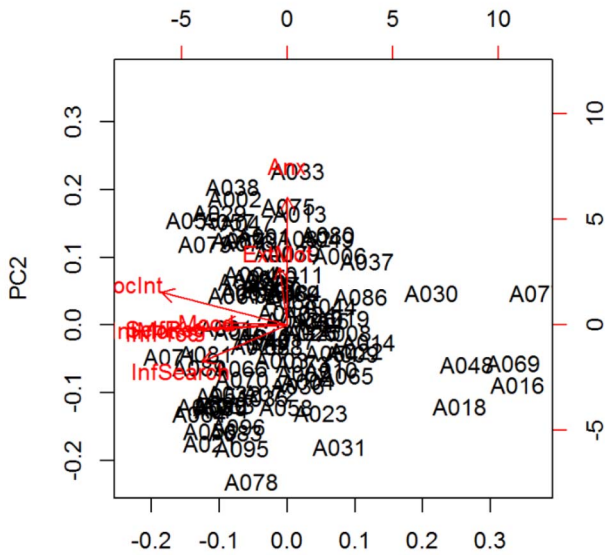


Figure 4. Biplot of the 8 dimensions for the student sample. Labels A01, A02, etc. identify individual students and their location on the PCA plane. Arrows show the specific dimension as given in Table I.

Each arrow in the diagram indicates the most representative direction of each dimension for the student sample. It is worth mentioning that in this figure, the smaller the angle between two arrows, the larger the correlation between the corresponding dimensions. Note that arrows are divided into two subgroups: 6 arrows point to the lower left region of the plot and are relatively packed, while the remaining 2 (Anx and ExtMot) are small and point upwards. These results suggest that there is certain correlation among the 6 dimensions of the first subgroup, and between the remaining 2.

2. Correlation Analysis (CA).

Correlation analysis is another method for studying the relationship that may exist among the 8 different dimensions used in this work.

The correlations among the dimensions are shown in Fig. 5. The larger a coefficient value is in the table, the stronger is the correlation between the corresponding dimensions.

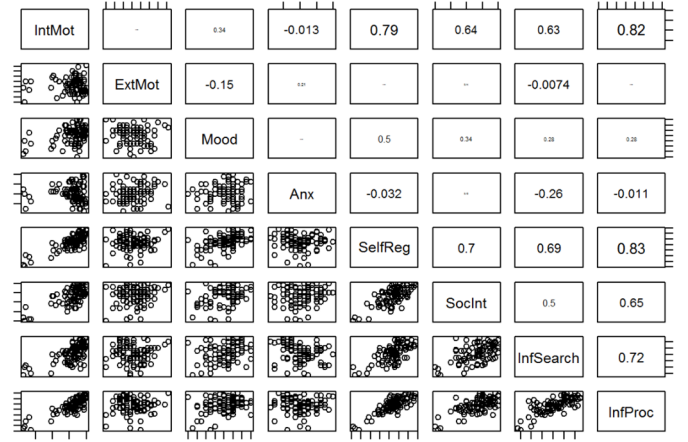


Figure 5. Correlation Matrix. Correlation coefficients between dimensions

From Fig. 5, it is seen that the highest correlation values are obtained between the following dimensions: *i)* self-regulation and strategies for the use and processing of information (0.83), *ii)* intrinsic motivation and strategies for the use and processing of information (0.82), and *iii)* intrinsic motivation and self-regulation (0.79).

3) Clustering

Using a hierarchical classification technique, together with the *Ward metrics* included within R, 3 clusters were formed including the 8 dimensions. The populations of the 3 resulting clusters were $N_1 = 72$, $N_2 = 6$, and $N_3 = 18$ (see Table V). The main features of these clusters were analyzed using bar and radar diagrams as explained below.

TABLE V. CLUSTER CHARACTERISTICS

	Cluster 1	Cluster 2	Cluster 3
<i>Size</i>	72	6	18
<i>Average Final Grade (AFG) (Scale: 0 - 100)</i>	73.6	75.5	81.9
<i>Number of students who failed the subject (N_F)</i>	20	1	3
<i>Standard Deviation</i>	13.3	10.9	12.4

In Fig. 6, the results for each cluster are represented by a set of 8 bars, corresponding to the dimensions listed in Table I. The bar height represents the average value for the corresponding dimension in a 1 to 5 scale. As observed in Table V, although Cluster 1 is the largest ($N_1 = 72$), it is the group with the lowest academic performance ($N_F = 20$ and $AVG = 73.6$). It has also the highest values of the dimensions of Anxiety and Extrinsic

Motivation among the three clusters; that is, the students of this group have mismanaged their anxiety and require relatively high extrinsic motivation. On the other hand, it is seen that Cluster 2 is characterized by low values for all dimensions. Finally, Cluster 3 shows an ideal profile for students, because they have high levels in all dimensions, except that they show the lowest dependency on extrinsic motivation and have the lowest levels of anxiety among the three clusters. Although its population ($N_3 = 18$) is smaller than that of Cluster 1 ($N_1 = 72$), it is the group with the best academic performance ($N_F = 3$ and $AVG = 81.9$). The characteristics for all clusters are best spotted in the radar diagram shown in Fig. 7.

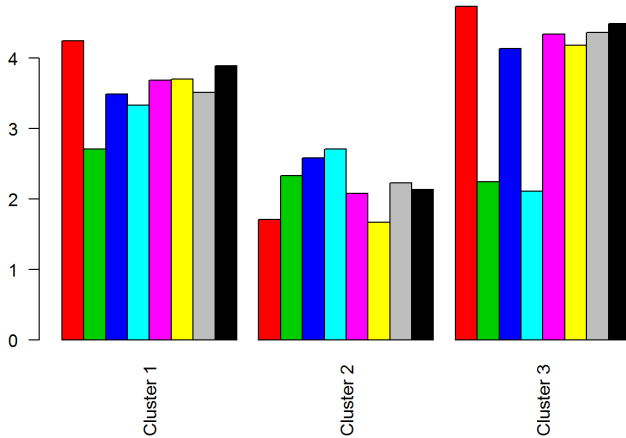


Figure 6. Bar Diagram for Cluster Comparison. The results for each cluster are represented by a set of 8 bars that correspond to their dimensions in the same order as listed in Table I. Each bar height represents the average value for a specific dimension in a 1 to 5 scale.

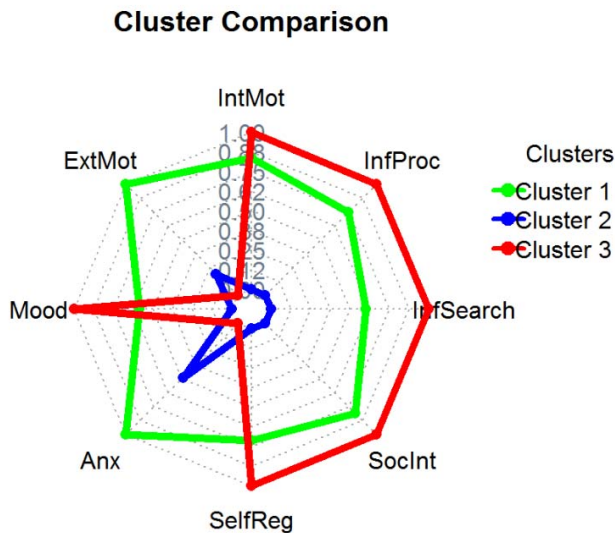


Figure 7. Radar diagram for cluster 1, 2 and 3. Results are parameterized in such a way that in each dimension a value of 1 is assigned to the cluster with the highest value and 0 to the lowest value, thus the remaining cluster has intermediate values.

G. Predictive Models

Taking into account the student profiles obtained previously, predictions on student performance were made by three

methods: a) determining grade averages in each cluster; b) classifying each student as *Pass* or *Fail* using decision trees, and c) identification of the dimensions that have an effect on the students' final grade by means of a regression analysis.

a) Prediction by clusters

The classification in the *Pass* or *Fail* categories was made with the Rattle library using R statistical package, taking as attributes the dimensions' values of the students.

Average final grades and the corresponding standard deviations for the 3 clusters mentioned above are presented in Table V above. The sample size was insufficient to algorithmically determine the appropriate number of clusters; therefore, it was decided subjectively to form only 3 clusters.

b) Classification using decision trees

The sample data table was divided selecting randomly 77 students to train the algorithm and 19 students to test it. The classification in *Pass* or *Fail* categories was made with the Rattle library within the R statistical package, taking as attributes the values of the dimensions of the students. A typical tree where the 8 dimensions were introduced is shown in Fig. 8.

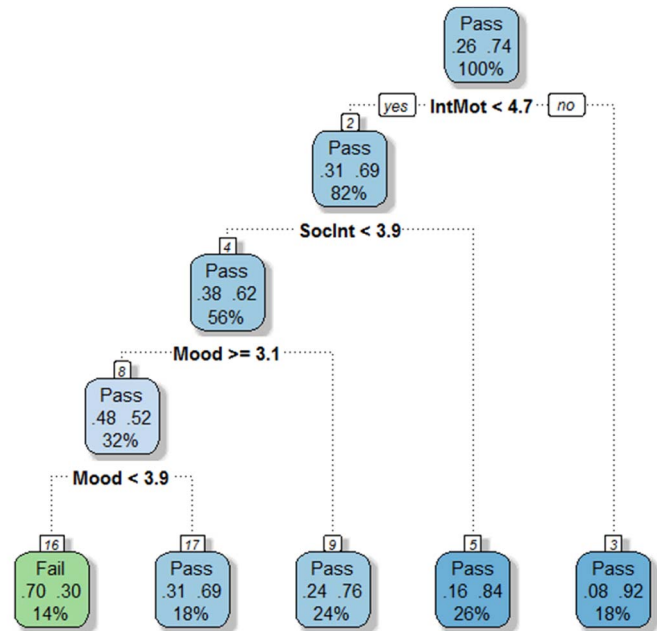


Figure 8. Tree diagram of classification rules

As it can be observed in Fig. 8, the Intrinsic Motivation dimension (*IntMot*) resulted to be the main variable to classify a student as *Pass* or *Fail*. The rule states that if the intrinsic motivation of a student is equal or greater than 4.7 (that is, very high), then the student will pass the course. Taking this rule as unique, 18% of the students are classified in an unambiguous way. At the other hand, if the Intrinsic Motivation is smaller than 4.7, the following variable to be considered is the Social Interaction dimension (*SocInt*), and the rule indicates that if the student has a social interaction value equal or greater than 3.9 then she will pass the course. With this second rule, 28% of the students are classified, and 84% of them are correctly classified while the remaining 16% are misclassified.

The last rule in the tree involves the Fitness and Mood dimension (*Mood*), in which the student will tend to fail if this dimension lies between 3.1 and 3.9. Therefore, to identify students who are more likely to fail than to pass the course, three basic rules have to be met: *i*) their intrinsic motivation value is smaller than 4.7; *ii*) their social interaction value is equal or greater than 3.9, and *iii*) their Fitness and Mood lies between 3.1 and 3.9.

TABLE VI. CONFUSIONS MATRIX USING THE COMPLETE STUDENT SAMPLE

		Prediction	
		<i>Pass</i>	<i>Fail</i>
Real	<i>Pass</i>	58	14
	<i>Fail</i>	15	9

The confusions matrix for the whole student sample ($N = 96$) is presented in Table VI. As can be seen, 70% of the predictions are correct, with 81% correct predictions for *Pass* and 38% correct predictions for *Fail*.

c) Regression analysis

The regression analysis considered the final student grade as the dependent variable in a 0 – 100 scale, and the 8 dimensions as the explanatory variables. The Stepwise method was chosen in order to make the best selection of the explanatory variables. In Tables VII and VIII, respectively, two Eviews are presented: the regression with all 8 dimensions included, and the best regression found taking into account only significant variables at level $p < 0.02$.

TABLE VII. EVIEW REPORT OF THE MULTIPLE REGRESSION USING ALL 8 DIMENSIONS AS VARIABLES IN ORDER TO EXPLAIN THE FINAL GRADE ^A

<i>Variable</i>	<i>Coeff.</i>	<i>p value</i>
<i>C</i>	98.16	0.0000
<i>INTMOT</i>	2.169	0.4874
<i>EXTMOT</i>	-2.986	0.1343
<i>MOOD</i>	-3.658	0.0872
<i>ANX</i>	-4.046	0.0254
<i>SELFREG</i>	-3.398	0.4659
<i>SOCINT</i>	1.428	0.5090
<i>INFSEARCH</i>	-1.565	0.6140
<i>INFPROC</i>	3.645	0.4040

A Eview report least square method with $N = 96$ observations, $R^2 = 0.1566$, S.E of Regression = 12.71, Sum Squared residual = 14068.76 and Durbin Watson Stat = 2.09

TABLE VIII. BEST REGRESSION INCLUDING ONLY THE MOST SIGNIFICANT VARIABLES^B

<i>Variable</i>	<i>Coeff.</i>	<i>p value</i>
<i>C</i>	96.478	0.0000
<i>MOOD</i>	-4.312	0.0203
<i>ANX</i>	-3.410	0.0358
<i>INTMOT</i>	3.013	0.0895
<i>EXTMOT</i>	-3.091	0.1085

B. Eview report least square method with $N = 96$ observations, $R^2 = 0.1434$, S.E of Regression = 12.53, Sum Squared residual = 14287.81 and Durbin Watson Stat = 2.1

From Table VII it is observed that the most significant variables are the Anxiety dimension, followed by the Fitness and mood dimension. However, it is well known that a regression including too many variables presents collinearity problems. Therefore, in the second regression (Table VIII), only the most significant variables were included.

V. ANALYSIS AND DISCUSSION

From the results obtained above, the first benefit is the fact that the students knew their profile described by the 8 dimensions of self-regulation, learning strategies and affective strategies. Knowing their areas of opportunity regarding each dimension, the students were motivated and agreed to participate in activities and strategies to address and expand these dimensions.

From the reliability study, 6 out of the 8 dimensions had Chronbach's alpha values greater than 0.700, except Anx and ExtMot. This could be due to their relatively low number of items (4 and 5, respectively). We consider that increasing the number of items of these dimensions could improve their statistical reliability.

Some of the data mining techniques mentioned above were also applied to individual sections, so that teachers could use the information of the student profiles in each section to establish opportunely the best teaching strategies and techniques of acquisition of skills for their students.

Regarding the Principal Component Analysis, we identified that 6 of the 8 dimension are grouped in one direction while Anxiety and Extrinsic Motivation point at a different direction. This result suggests that there is a certain correlation among the 6 dimensions of the first subgroup, and that they are relatively independent from the remaining 2.

From the correlation matrix it can be seen that the highest correlations are among the following pairs of dimensions: Self-regulation and Strategies for the use and information processing (0.83), Intrinsic Motivation and Strategies for the use and information processing (0.82), and Intrinsic motivation and Self-regulation (0.79). This suggests that students which are intrinsically motivated and that show appropriate self-regulation skills are best suited to select and use better strategies from information processing.

In the clustering process, the sample was divided into 3 groups, with populations, $N_1 = 72$, $N_2 = 6$, and $N_3 = 18$. Group number 1 is the largest and has the lowest average and more students who failed the course. We found that this group has higher values for anxiety and requires more extrinsic motivation (Fig. 7). Group 2 is the smallest one, and has the less developed dimensions, except Anxiety and Extrinsic motivation. Finally, group 3 is the one with the highest final average grade, few failed students and has all dimensions highly developed, except Anxiety and Extrinsic motivation. This results suggest that high levels of anxiety and the need of extrinsic motivation may adversely affect student academic performance.

Our decision tree shows that if the intrinsic motivation is very high, students are very likely to pass the course. The second important variable corresponds to social interaction. If it is greater than or equal to 3.9, although intrinsic motivation is not so high, students have a good chance of passing. If fitness and mood is less than 3.1, given the above condition, the student has also a good chance of passing. If fitness and mood is greater than or equal to 3.9, the student is also likely to pass. Finally, if the intrinsic motivation is less than 4.7, social interaction is less than 3.9 and their fitness and mood is between 3.1 and 3.9, the student has high likelihood of failing the course. The most important variables that explain the possibility that the student passes the course are therefore intrinsic motivation, social interaction and fitness and mood.

On the other hand, from the confusion matrix it is found a 70% of correct predictions of students that approved or not the course. It was also obtained a 81% of correct predictions of students approving the course, and a 38% of correct predictions of students failing the course.

Finally, from the multiple regression analysis, it is found that the most significant variables explaining the final student average grade are Fitness and mood, Anxiety, Intrinsic motivation and to some extent Extrinsic motivation.

VI. CONCLUSIONS AND FUTURE WORK

An instrument for assessing self-regulation skills, affective strategies and learning of engineering students, grouped in 8 dimensions was adapted. These 8 dimensions are: *i)* intrinsic motivation, *ii)* extrinsic motivation, *iii)* fitness and mood, *iv)* anxiety, *v)* self-regulation, *vi)* social interaction, *vii)* searching and selection of information strategies, and *viii)* processing and use of information strategies. The statistically reliability of the used instrument was tested through a Cronbach's alphas approach. It was determined that, in general, the dimensions are properly characterized and differentiated, enabling us to determine *student profiles*. However, we consider that the Anxiety and Extrinsic motivation dimensions should be explored in more detail to increase their Cronbach's alpha values.

Our *social validation* shows very good agreement between the profiles assigned by the instrument and those perceived by the students. Indeed, it was found that 95% of students agree

with their assigned profiles. For stronger validation, in a future I would try to do multi-campus and multi-institutional studies.

Statistical techniques and data mining applied properly to the data allowed us to group students in different clusters according to their student profiles.

We are currently working on automating data mining functions to obtain timely input information for the professor in terms of their students' profiles and results. Future work also includes considering multiple intelligences, learning styles and diagnostic tests about previous knowledge, for larger student samples. Our working hypothesis is that this information can be used to guide the teacher in order to adopt appropriate teaching strategies aimed to improve overall student academic performance.

Although the main limitation of this study is the relatively small sample size of $N = 96$ students, our results encourage us to continue and extend this study to strengthen the findings. As future agenda, the introduction of new variables like the student past academic records and other student related meta-data might be included in the analysis. This will allow us to improve predicting student outcomes and to alleviate the course drop-rate by means of more student-tailored actions. We are currently working on expanding the sample of students. The ultimate goal of the project in the future will be to set the path in order to design and build a predictive model to better understand student academic behavior.

ACKNOWLEDGMENT

We would like to thank Roberto Alejandro Cárdenas-Ovando and David Escobar-Castillejo, PhD students, for their contributions to implement this system, as well as Héctor Alberto Rueda-Zárate for helping in the integration of the database and data processing with R.

REFERENCES

- [1] J. Noguez, O. Olmos, L. Neri, A. González-Nucamendi, G. Aguilar, G. Alanís, V. Robledo-Rella, M.C. Romero, M.A., Hernández, and G. Cervantes, "Final Report on Learning Analytics for Adaptive Learning" [Reporte Final Analíticas para el aprendizaje adaptativo]. Rectoría Zona Centro Sur. Tecnológico de Monterrey, Mexico City, 2014.
- [2] V. Robledo-Rella, H. Rueda, and R.A., Cárdenas, "Characterization of student learning styles at the Tecnológico de Monterrey using Felder-Silverman ILS", 2do Congreso Internacional de Innovación Educativa, Tecnológico de Monterrey, 2015.
- [3] J. Noguez, D.A. Escárcega, D.E. Castillejos, "Validation of instrument for multiples intelligences and learning strategies". [Validación de instrumento para inteligencias múltiples y estrategias de aprendizaje], 2do Congreso Internacional de Innovación Educativa, Tecnológico de Monterrey, 2015.
- [4] L. Neri, J. Noguez, and G. Alanís-Funes, "Validation of an instrument for determining students' self-regulation skills, affective strategies and learning strategies". [Validación de un instrumento para determinar habilidades de autorregulación, estrategias afectivas y de aprendizaje de los alumnos]. 2do Congreso Internacional de Innovación Educativa, Tecnológico de Monterrey, 2015.
- [5] Gargallo B., Suárez-Rodríguez J. M., Pérez-Perez, C. (2009). El cuestionario CEVEAPEU. "An instrument for assessing the learning strategies of college students". [Un instrumento para la evaluación de las estrategias de aprendizaje de los estudiantes universitarios]. RELIEVE, v. 15, n. 2, p. 1-31.

- [6] B.J., Zimmerman, and D.H. Schunk (Eds). "Self-regulated learning and academic achievement: Theoretical perspectives". 2a Ed., Mahwah, NJ: Erlbaum, 2001.
- [7] D.H., Schunk (2012). "Learning theories. An educational perspective". 6th ed. Pearson Education.
- [8] Fernández, L. M. (2009). "Adaptation of two motivation questionnaires: learning self-regulation and learning environment". [Adaptación de dos cuestionarios de motivación: Autorregulación del Aprendizaje y Clima de Aprendizaje]. *Persona*, (12), 167-185.
- [9] Caso J., Vargas, L.P., Contreras, L.A., Rodríguez J.C., Urias E. (2010). "Psicometric properties of the academic self-regulation questionnaire" [Propiedades psicométricas del Cuestionario de Autorregulación Académica]. UEE. Reporte Técnico Num. 10 - 003. Ensenada, México. Universidad Autónoma de Baja California.
- [10] M.A. Arriola, "Relationship between learning strategies and self-regulation. An explanatory model". [Relación entre estrategias de aprendizaje y autorregulación: Un modelo explicativo]. PhD thesis. Universidad Iberoamericana, México D.F. México, 2001.
- [11] Boone, H. N. & Boone, D. A. (2012). Analyzing likert data. *Journal of extension*, 50(2), 1-5
- [12] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-33