JMETP
Journal for the Mathematics Education and Teaching Practices, 1(2), 63-72, Dec 2020
e-ISSN: 2717-8587
dergipark.org.tr/jmetp

Young Wise Publishing
youngwisepub.com
© 2020

Research Article

Developing a risk model to identify factors which critically affect secondary school students' performance in mathematics

Stylianos Tsakirtzis[1] and Ioannis Georgakopoulos [2]

Department of Industrial Design and Production, University of West Attica, Greece

Article Info

Received: 04 October 2020
Revised: 17 November 2020
Accepted: 10 December 2020
Available online: 15 Dec 2020

Keywords:
Risk factors
Risk model
Students' achievement
Students' engagement

Abstract

A concrete knowledge on Mathematics is essential on the ground that it constitutes to be a key-ingredient to a brilliant academic career. Though, a lot of students encounter insurmountable difficulties and as a consequence they fail their Mathematical courses. That holds true particularly on the case of secondary school students. Thereby, controlling the risk of students' failure in Mathematics is of utmost importance. The paper demonstrates a risk model which identifies factors that critically affect secondary school students' performance and prioritize them according to their contribution to the risk occurrence. The risk model has been built on the base of a binary logistics regression analysis on students' behavioral engagement data. These data reflect students' effort and involvement in the entire learning process. The risk model development process is presented in the context of a case study on a specific Mathematical course, delivered at a Greek private Secondary School (Gymnasium). The binary logistics' regression outcome has proved that students' achievement on schoolwork and review packages are factors which critically affect the students' performance in the respective course. It is also important to highlight that schoolwork completed appeared to have significant contribution to the risk occurrence, indicating that schoolwork completed could be regarded as a cardinal factor which critically affects students' performance in the context of the respective study.

To cite article

Tsakirtzis, S., & Georgakopoulos, I. (2020). Developing a risk model to identify factors which critically affect secondary school students' performance in mathematics. *Journal for the Mathematics Education and Teaching Practices, 1*(2), 63-72.

Introduction

The teaching of Mathematics aims at achieving the organization of students' thoughts, which is essential for students' overall cultivation. For that reason, modern teaching is tailored to students' needs with a view to cultivating their different abilities and skills (Tomlinson, 2001, 2003).

However, emphasis should be placed on the way a Mathematical course is delivered at Greek Secondary School (Gymnasium). It is vital to highlight that the curriculum and syllabus for the teaching of Mathematics at Gymnasium stress on the fact that the teaching of Mathematics should aim at practicing students' rational and disciplined thinking and connecting Mathematics to the real world (Instructions on the teaching of Mathematics for the years, 2018-2020). In line with the previously cited objectives, teachers resort to a plethora of instructional methods. These methods may vary in some aspects but they share a specific characteristic (Carolan & Guinn, 2007). They view knowledge as 'student property' which is gained when a student is actively involved in the learning process. Thereby, students should be urged to actively participate in the activities performed during class (Sullivan et al. 2011). Hence, an important metric in the territory of students' active participation is the students' schoolwork.

Another crucial issue on the didactic of Mathematics is the students' evaluation. The assessment of students' final achievement is based on the below pillars (Tzoka, 2019):

- The degree of achievement of learning objectives

[1] Corresponding Author: Dr. Stylianos Tsakirtzis, Department of Industrial Design and Production, University of West Attica, Greece. E-mail: s_tsakirtzis@hotmail.com Orcid ID: 0000-0001-5856-2098
[2] Dr., Ioannis Georgakopoulos, Faculty of Industrial Design Engineering, and Accounting and Finance., University of West Attica, Greece E-mail: igtei@uniwa.gr Orcid ID: 0000-0003-4311-350X

- The appropriateness of the methodology used
- The degree of student-teacher cooperation achieved
- The proper use of many different sources of information
- The development of synthetic and creative capacity as well as the development of self-energy
- The interest shown by the students.

According to Tzoka (2019), students could better comprehend and assimilate the mathematical concepts and ideas through their involvement in mathematically challenging tasks in class. Students who usually fail a mathematical course are individuals who very easily lose interest in learning. Therefore, such students should constantly be involved into activities that require hard work so that they continue maintaining a good extent of interest in learning. In line with the respective curriculum and syllabus, the underlying Private Secondary School uses a specific evaluation method. Some activities that are assessed are listed below:

- The overall participation of a student in the learning process (the questions he asks; the answers he gives; his contribution to the study of a subject in the classroom; his cooperation with classmates; the diligence in the execution of the tasks assigned to him; his comprehension on concepts and phenomena; his problem-solving and communication skills and his critical thinking and creativity);
- The daily tasks performed by a student at school or at home;
- Individual or group synthetic, creative and interdisciplinary work;
- Hourly/few-minute written tests;
- Baccalaureate exams.

It is also essential to highlight that the grade reflecting students' annual performance (graduation grade) is computed by the grades of the first quarter, the second semester and the written summary examination.

Our research interest is directed into students' behavioral engagement on the ground that this can be measured by graded activities in accordance with the evaluation method used at the respective Private Secondary School. Given that students' behavioral engagement is reflected on their effort (Hopf et al. 2003), our research question could be shaped as follows:

Does students' behavioral engagement, reflecting on their effort (exercises completed, assignments completed, schoolwork completed), critically affect their achievement? It is important to highlight that the term critical achievement insinuates a threshold below which students at risk could be identified.

In order to answer that question, we have developed a risk model through a binary logistics regression on students' behavioral engagement data. These data are candidate risk factors of students' failure. The risk model decides which of these data have real contribution to the risk occurrence of students' failure by indicating which of the respective data are statistically significant. In addition, the risk model prioritizes the risk factors according to their contribution to the risk occurrence pointing out the contribution of each factor to the reduction in the probability of risk occurrence.

The risk model development is demonstrated through a case study on a specific Mathematical course delivered at a Greek Private Secondary School (Gymnasium). The next sections shed more light on the way the risk model has been built and the findings of our research.

Literature Review

Factors Affecting Secondary School Students' Final Achievement in Mathematics

In the generic territory of secondary school students' achievement there are many researches that have proved the association of secondary school students' learning outcome with their engagement (Frederick et al. 2004; Marks, 2000; Willms, 2003). In parallel manner, researches have proved the correlation of secondary school students' achievement to their effort during class (Hopf et al. 2003). Another essential factor affecting secondary school students' achievement in Mathematics is self-efficacy (McConney & Perry, 2010; Yurt, 2014). Additionally, another study points out that high school students' achievement is dependent on psychological; behavioral and demographic factors (Casillas et al. 2012). Needless to say, that the behavioral factors reported are affiliated with students' interaction with the learning activities and the entire learning process such as the homework completed and the study time. Finally, the attitude of secondary school students (middle school students and high school students) towards Mathematics has also been reported as an essential factor having significant impact on their performance (Hemmings et al. 2011).

Secondary School Students at Risk in Mathematics

A study (Flores & Kaylor, 2007) has indicated that secondary school students at risk in mathematics could be predicted through a proper analysis of curriculum-based data. The respective study has made use of a t-test method to assess students' progress in the context of pre and post curriculum based tests, reporting that curriculum based data could be strong predictors of students' critical performance. Other studies (Kajander et al. 2008; Xin et al. 2005) have underlined the significant role of teaching method to prevent secondary school students from failing a mathematical course. The study of (Kajander et al. 2008) also stresses on the need for early intervention for students at risk. Additionally, a multiple regression analysis on students' engagement data has been used in another study (Sciarra & Seirup, 2018) and the analysis' outcome has proved that the cognitive and behavioral engagement are stronger predictors of secondary school students' critical achievement (students at risk) in Mathematics in comparison to the emotional engagement.

Summing up, there is not a specific set of factors which critically affect secondary school students' final achievement in Mathematics that has been reported in literature. However, in the light of some studies, students' engagement data appear to be a significant factor.

Thereby, its' not easy to predict secondary school students' performance in Mathematics given that the factors which critically affect their final achievement is course-oriented. Though, prediction models could be generated on the base of the identified factors for specific courses. It is important to highlight that the researches which have been previously referred to, state a set of factors which affect students' performance in the respective courses without having their results appertained to a specific risk management framework. For that reason, our risk model which is demonstrated in this paper is based on a concrete risk management framework, the application of which could lead to a potent prediction model and an impending warning system. Nevertheless, a warning system generation process denotes that the prediction model should be verified in terms of a plethora of similar courses sharing the sane learning design. Therefore, our risk model could be verified in the context of many similar Mathematical courses in order to come up with a proper prediction model on the base of which a warning system could be developed. In parallel manner, the warning system will achieve the control of students at risk and could be delivered to educators. It is essential to highlight that students at risk in Mathematics should be identified and controlled on the ground that a lot of students fail their Mathematical courses. The way our risk model has been built in the light of the respective framework is presented in the next section.

Method

Research Model

According to Vose (2008), the risk factors' identification process is part of a risk analysis process which is also part of a generic risk management framework. The risk factors' identification process aims at developing a risk model to identify factors which have significant contribution to the risk occurrence and prioritize them analogically to their contribution. The identified risk factors could constitute the base on which a prediction model could be generated and the verified prediction model could also set the standards for the development of a warning system.

In our research, we have developed two risk models, one developed on the graduation base of 10 and one developed on the numeric threshold of 12 in the context of students' critical performance. Our work is based on a specific methodology used to identify risk factors of students' failure in e-learning courses (Georgakopoulos et. al. 2018). Nevertheless, the respective methodology has not been tested on courses delivered in a conventional way. Thereby, our research attempts to expand the underlying methodology to cover the needs of conventional teaching. The respective methodology includes the below phases:

- Data Collection
- Risk Model Development
- Prediction Model Generation
- Prediction Model Verification
- Warning System Development

It is important to clarify that only the phases 1 and 2 are demonstrated in this paper on the ground that the paper takes up the issue of risk factors' identification and doesn't take up the issue of students' final achievement prediction.

Data Collection

Getting perspective on the data collection process, it is important to explain that the data collected in terms of the underlying methodology in the work referred (Georgakopoulos et al. 2018), were data related to students' interaction with Moodle LMS. Such data included students' logins into the system, students' completed activities, time spent on system, time allotted to activities, resources' view and posts on forum. Though, it is essential to clarify that these data are typically stored into Moodle LMS and thereby it was easy to gather them. The data were collected after the first run of the course.

However, a conventional course is not always connected to a LMS and on this account it is not easy to collect the students' behavioral engagement data. To answer that purpose, in our case, we collected the requisite data from all graded activities in the framework of the evaluation method used at the respective Private Secondary School.

In the spirit of the above activities, we gathered the collective students' engagement data out of three grades (Grade A; Grade B and Grade C) in respect to a specific mathematical course delivered at the private Secondary School. The data were collected from an official school database within a three years period. Needless to say that 152 students were enrolled into the course within that specific period of time and the data were collected out of all enrolled students. It is important to stress on the fact that conventional teaching, reflecting on lectures, schoolwork, homework and exercises was included in the course delivery process. No part of the course was mounted on a learning management system. Thereby, each engagement data item was measured by the graded activities on the ground that students' study (completion of study material) could not easily be assessed in a conventional course delivery mode.

In a more elaborate detail, the data were collected by the students record held at the respective Private Secondary School. Given that the data should be irrefutable, we chose to use only those concerning the last three (3) years. It is important to stress on the fact that the registration process should have been completed in three years' time. In the registration process, teachers register students' scores in the official database of the Ministry of Education in our country (My School Database). The data set collected is listed into table 1. The table 1 also shows when each of the specific data item was measured:

Table 1.
Data Collected

Data	Measured (Time period)
Average Theory Test Score	Weekly
Percentage of Exercises Completed	Daily
Average Homework Score	Weekly
Percentage of Review Packages Completed	Monthly
Average Review Packages Score	Monthly
Percentage of School Work Completed	Daily
A' Semester Rating	At the end of A' semester
B' Semester Rating	At the end of B' semester
Absences	Daily
Score of Final Examination	After the final exams
Graduation Grade	Annually

It is important to come up with some vital clarification in regard to the data collected:

Average Theory Test Score: It is computed as the average of the grades in the written tests performed during each semester.

Percentage of Exercises Completed: It is the percentage of exercises completed by each student assigned by teachers as daily homework. An exercise is deemed to be completed by a student only whether a student has answered all the questions included in the exercise.

Average Homework Score: It is computed as the average evaluation score out of all students' homework assignments.

Percentage of Review Packages Completed: It is the percentage of exercises completed by each student in terms of the review packages. Such packages are given to students when a didactic unit is completed as well as during the Christmas and Easter holidays. A review package is viewed as completed only on the case that all exercises included are completed.

Average Review Packages Score: It is computed as the average evaluation score out of all students' review packages.

Percentage of School Work Completed: It refers to the exercises that students are asked to do in class during a lecture. More specifically, the teacher delivers the course (Theory, Exercise, Methodologies, Examples) and then the students are invited to elaborate on theory and exercises. In that way, students assume an active role in the learning process.

Risk Model Development

Along with the underlined data listed in table 1, we modeled the binary variable srisk as the variable describing students who failed the course. The state '0' was modeled to describe students who passed the course whereas the state '1' was modeled to describe students who fell through (Macfayden & Dawyson, 2010). Typically, these are the states used in some researches which are related to students at risk in e-learning courses (Macfayden & Dawyson, 2010; Georgakopoulos et al. 2018; Anagnostopoulos et al. 2020). The state "0" in the studies referred, was modeled to describe students who passed the course denoting that risk is not occurred in that state. In parallel manner, the state "1" was modeled to describe students who failed the course implying that risk is occurred in that state. Needless to say that the graduation grade defined the numeric threshold on the base of which students at risk were identified. All the variables which were modeled to describe the respective data (variables modeled to describe the engagement data along with the variable modeled to describe students at risk) are listed into table 2. The first column on table 2 points out the data collected and the second column indicates the variable's name which is given to each variable modeled.

Table 2.
Variables Modeled

Data Description	Variable Modeled
Average Theory Test Score	thg
Percentage of Exercises Completed	exc
Average Homework Score	hg
Percentage of Review Packages Completed	pc
Average Review Packages Score	pg
Percentage of School Work Completed	swc
A' Semester Rating	p1g
B' Semester Rating	p2g
Absences	tabs
Score of Final Examination	grade
Graduation Grade	gradeg
Students at risk	srisk

At first, we modeled the variable srisk on the base of the numeric graduation grade threshold of 10. We deployed this data set in terms of a binary logistics regression analysis (Macfayden & Dawyson, 2010; Georgakopoulos et al. 2018) after the first course run and we came up with an initial model to identify which of the variables modeled could be regarded as risk factors having significant contribution to the risk occurrence. Needless to say that "srisk" was the dependent variable and the other variables were the independent variables (coefficients). The "gradeg" variable, describing the graduation grade, was not entered into the coefficients on the ground that it only defined the variable "srisk". In parallel manner, the variables "grade", "p1g", "p2g", reflecting the final examination score and the A' and B' semester rating respectively, were not entered into the coefficients given that they only contributed to the graduation grade. It is also essential to explain that all independent variables which were entered into the coefficients ("hc", "swc", "pc", "pg", "thg", "hg", "tabs") were measured as Scale, whereas the dependent variable "srisk" was measured as Nominal. Additionally, we moved on increasing that threshold in order to examine the liability of common risk factors. Therefore, increasing the threshold by 2 units (12 instead of 10), as proposed by researches (Hopf et al. 2003) and executing the same process (binary logistics regression analysis) we came up with another model.

The binary logistics regression in some studies related to students at risk (Macfayden & Dawyson, 2010; Georgakopoulos et al. 2018) has been used to achieve the below goals:

- Identify factors which critically affects the students' performance and have contribution to the risk occurrence;
- generate a prediction model for students at risk.

Our models do not only identify the risk factors of students' failure but they also classify students into two groups (students at risk; students not at risk). Nevertheless, emphasis is not laid on the prediction aspect of our models on the ground that this paper is encircled on the identification of factors which critically affects the students' final achievement.

Results

The binary logistics regression outcome (on the base of threshold 10) has led to a specific risk model. The table 3 sheds light on some cardinal performance characteristics of our model.

Table 3.
Performance Characteristics (Risk model-threshold 10)

Performance metrics	Value
Sensitivity	0.857
Specificity	0.957
Precision	0.857

The table 3 points out that our model could be deemed to be a good model given that achieves high score in each performance metrics territory (Sensitivity: 85.7%; Specificity:95.7%; Precision:85.7%). Emphasis should be placed on the precision metric which reflects the intended classification percentage, denoting that in our case, the model achieves an 85.7 classification percentage. Looking at table 3, we can deduct that the intended classification outcome (precision) matches with the real classification outcome (sensitivity). The high specificity percentage denotes that the majority of students at risk are correctly classified. However, the same doesn't hold true for students not at risk. Thereby, our model classifies correctly the 85.7% of the cases insinuating that a small portion of cases (14.3%) is not correctly classified. Thus, some students who should have been classified into the group of students not at risk have finally been classified into the other group (students at risk). However, the high classification percentage which this model achieves indicates that the amount of students who have not been correctly classified is not significant.

In parallel manner, it is important to underline that our model accounts for the 78.5 % of the liable risk factors (Nagelkerke R^2) implying that approximately only the 21.5 % of the liable risk factors is not identified (see table 4). It is important to stress on the fact that the range for Nagelkerke R^2 is between 0 and 1. The value "1" denotes a perfect model fit (Allison, 2014; Smith et al., 2013; Hair et al. 2006). On the ground that the Nagelkerke R^2 value for our model is close to 1, our model could be deemed to be a good model.

Table 4.
Model Summary (Risk model-threshold 10)

Model Summary

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	164.037	166.037	169.060	151						
H_1	53.141	73.141	103.380	142	110.895	< .001	0.676	0.785	0.126	0.518

The table 5 shows the coefficients which could be included in the regression model according to the p-value.

Table 5.

Coefficients (Risk Model-threshold 10)

Coefficients

					Wald Test		
	Estimate	**Standard Error**	**Odds Ratio**	**z**	**Wald Statistic**	**df**	**p**
(Intercept)	9.363	3.716	11654.011	2.520	6.351	1	0.012
hc	-6.044	5.418	0.002	-1.116	1.244	1	0.265
swc	-13.734	6.022	1.085e -6	-2.281	5.202	1	0.023
pg	-0.730	0.345	0.482	-2.118	4.484	1	0.034
thg	0.145	0.125	1.156	1.154	1.332	1	0.248
hg	0.064	0.356	1.067	0.181	0.033	1	0.856
tabs	0.026	0.119	1.027	0.222	0.049	1	0.825
pc	6.383	4.216	591.757	1.514	2.292	1	0.130

Note. srisklr level '1' coded as class 1.

The factors which have statistically significant contribution to the occurrence of the risk of students' failure are derived from coefficients with p-value lower or equal to 0.05. Thereby, according to table 5, in our case, these factors are the percentage of School Work Completed and Average Review Packages Score. Therefore, our regression model could be given as follows:

Logit(srisk)= -13.374* swc – 0.730*pg + 9.363

Looking at the estimates on table 5, we can deduct that an increase in the percentage of school work completed leads to a significant decrease (13.734) in the logarithm of the probability of risk occurrence. In parallel manner, the increase in the average review package score leads to a slight decrease (0.730) in the logarithm of the probability of the risk occurrence. Thereby, school work completed is a factor which has greater contribution to the risk occurrence in comparison to the review packages score.

Looking at the odds ratio on table 5, we can conclude that students who achieve a great percentage of schoolwork completed are 3.05 times as likely to turnout to pass than students who don't achieve a great percentage of schoolwork completed. In parallel manner, students who achieve a great score on packages are 0.482 times as likely to turnout to pass than students who don't achieve a great score on packages.

The binary logistics regression outcome (based on threshold 12) has led to another risk model. The table 6 sheds light on some cardinal performance characteristics of our model.

Table 6.

Performance Characteristics (Risk model-threshold 12)

Performance metrics	**Value**
Sensitivity	0.857
Specificity	0.957
Precision	0.857

The table 6 points out that our model could be deemed to be a good model given that achieves high score in each performance metrics territory (Sensitivity: 85.7%; Specificity:95.7%; Precision:85.7%). Looking at the table 6, we can deduct that the intended classification outcome (precision) matches with the real classification outcome (sensitivity). The high specificity percentage denotes that the majority of students at risk are correctly classified. However, the same doesn't hold true for students not at risk. Thereby, our model classifies correctly the 85.7% of the cases insinuating that a small portion of cases (14.3%) is not correctly classified. Thus, some students who should have been classified into the group of students not at risk have finally been classified into the other group (students at risk). However, the high classification percentage which this model achieves indicates that the amount of students who have not been correctly classified is not significant.

In parallel manner it is important to underline that our model accounts for the 78.4 % of the liable risk factors (Nagelkerke R^2) implying that approximately only the 21.6 % of the liable risk factors is not identified (see table 7). Needless to say that the range for Nagelkerke R^2 is between 0 and 1. The value "1" denotes a perfect model fit (Allison, 2014; Smith et al. 2013; Hair et al. 2006). On the ground that the Nagelkerke R^2 value for our model is close to 1, our model could be deemed to be a good model.

Table 7.

Model Summary (Risk Model-threshold 12)

Model summary

Model	Deviance	AIC	BIC	df	X^2	p	McFadden R^2	Nagelkerke R^2	Tjur R^2	Cox & Snell R^2
H_0	164.037	166.037	169.060	151						
H_1	53.310	69.310	93.501	144	110.726	< .001	0.675	0.784	0.122	0.517

The table 8 shows the coefficients which could be included in the regression model according to the p-value:

Table 8.

Coefficients (Risk Model-threshold 12)

Coefficients					**Wald Test**			**95% Confidence interval**	
	Estimate	Standard Error	Odds Ratio	z	Wald Statistic	df	p	Lower bound	Upper bound
(Intercept)	9.780	3.537	17677.009	2.765	7.645	1	0.006	2.848	16.713
hc	-6.072	5.426	0.002	-1.119	1.253	1	0.263	-16.706	4.562
pc	6.493	4.266	660.637	1.522	2.317	1	0.128	-1.868	14.854
swc	-13.468	5.925	1.415e -6	-2.273	5.166	1	0.023	-25.082	-1.855
tabs	0.034	0.117	1.035	0.293	0.086	1	0.769	-0.195	0.264
pg	-0.742	0.343	0.476	-2.163	4.680	1	0.031	-1.415	-0.070
hg	0.076	0.350	1.079	0.216	0.047	1	0.829	-0.611	0.762
thg	0.137	0.122	1.147	1.126	1.268	1	0.260	-0.102	0.376

Note. srisklr level '1' coded as class 1

The factors which have statistically significant contribution to the occurrence of the risk of students' failure are derived from coefficients with p-value lower or equal to 0.05. Thereby, according to table 8, in our case, these factors are the percentage of School Work Completed and Average Review Packages Score. Therefore, our regression model could be given as follows:

Logit(srisk)= -13.468* swc – 0.742*pg + 9.780

Looking at the estimates on table 8, we can deduct that the increase in the percentage of school work completed leads to a significant decrease (13.468) in the logarithm of the probability of risk occurrence. In parallel manner, the increase in the average review packages score leads to a slight decrease (0.742) in the logarithm of the probability of the risk occurrence. Thereby, school work completed is a factor which has greater contribution to the risk occurrence in comparison to the review packages score.

Looking at the odds ratio on table 8, we can conclude that students who achieve a great percentage of schoolwork completed are 2.05 times as likely to turnout to pass than students who don't achieve a great percentage of schoolwork completed. In parallel manner, students who achieve a great score on packages are 0.476 times as likely to turnout to pass than students who don't achieve a great score on packages.

Conclusions

Both risk models achieve a great score in each performance metrics territory (see table 3 and table 6). Both risk models account for a good percentage of the identified factors (see table 4 and table 7). The regression outcome for both models have proved that schoolwork completed and average review package scores are significant factors which have contribution to the risk of secondary school students' failure in the respective Mathematical course (see table 5 and table 8). As shown in these tables, the schoolwork completed has greater contribution to the risk occurrence in comparison to the average review package score. Hence, the schoolwork completed could be deemed to be a major risk factor which critically affects students' performance in that specific course. It is important to stress on the fact that a couple of researches have proved that schoolwork has great impact on the Mathematical learning process (Hopf et al. 2003). It is also essential to state that our research question has been verified on the ground that the percentage of school work completed which has been proved to be a significant risk factor, reflects students' effort and could be regarded as an indicator of the students' behavioral engagement.

Though, our sample is not sufficient enough to state that the schoolwork completed or the Average Review Packages score critically affect the students' achievement in all Mathematical courses on the ground that the risk factors of students' failure could vary among courses. However, our model could potentially be applied to a plethora of courses having the same learning design in order to come up with a risk model suitable for many similar courses.

Though, our sample is not sufficient enough to state that the schoolwork completed or the Average Review Packages score critically affect the students' achievement in all Mathematical courses on the ground that the risk factors of students' failure could vary among courses. However, our model could potentially be applied to a plethora of courses having the same learning design in order to come up with a risk model suitable for many similar courses. We are currently working on verifying our risk model in the context of many similar courses and thereby a prediction model is in the pipeline.

Recommendations

Our study has proved that risk factors which affect secondary school students' performance in Mathematics could be traced into students' behavioral engagement. Though, behavioral engagement is not the only aspect of students' engagement. The emotional engagement, which reflects students' attitude towards a specific course is also a cardinal part of students' engagement. In parallel manner, the students' emotional engagement could affect students' effort and thereby students' emotional engagement could be combined with students' behavioral engagement in an attempt to identify factors which critically affect students' performance in Mathematics. However, the emotional engagement cannot be easily measured. Therefore, we believe that there is much space for more researches in that territory.

Acknowledgment

We thank Mrs. Aikaterini Karahaliou, (Director and Owner) and Mrs. Aggeliki Ntouska (the head of the Private High School's IT department) for their help in the data collection process.

Biodata of Authors

Stylianos Tsakirtzis holds a BSc degree in Mathematics (University of Ioannina), an Msc degree in Applied mathematics – Mechanics (National Technical University of Athens) and a PhD degree in Non–Linear Dynamic (National Technical University of Athens). His research interests are in Non Linear Dynamics and also in student learning in mathematics, research and development of media learning, mathematical disposition and mathematics learning for Secondary school. He has participated in several projects as a research associate at the Piraeus University of Applied Sciences. Currently, Dr Tsakirtzis is an associate at the University of West Attica. **Affiliation:** University of West Attica, Faculty of Industrial Design Engineering, Greece. **E-mail**: s_tsakirtzis@hotmail.com **Scopus ID:** 8419225900

Ioannis Georgakopoulos holds a BSc degree in Industrial Informatics Engineering, (Piraeus University of Applied Sciences), an MSc degree in Educational Technology and Human Resources Development, (University of Athens) and a PhD degree in risk analysis techniques in e-learning, (University of Athens). His research interests are in e-Learning, risk analysis, learning management systems and learning analytics. Additionally, he has participated in several projects as a research associate at the Piraeus University of Applied Sciences. Currently, Dr Georgakopoulos is an associate at the University of West Attica. **Affiliation:** University of West Attica, Faculties: Industrial Design Engineering, Accounting and Finance, Greece **E-mail**: igtei@uniwa.gr

References

Allison, P. D. (2014). Measures of fit for logistic regression, Paper No. 1485-2014. Paper presented at the *SAS Global Forum 2014 Conference*, Washington D.C. 23-26 March, 2014.

Anagnostopoulos, T., Kytagias, C., Xanthopoulos, T., Georgakopoulos, I., Psaromiligkos, I., Salmon, I. (2020). Intelligent Predictive Analytics for Identifying Students at Risk of Failure in Moodle Courses. *International Conference on Intelligent Tutoring Systems. June, 2020, 152-162, Springer, Cham*

Carolan, J. & Guinn, A. (2007). Differentiation. *Lessons from Master Teachers. Educational Leadership*, *64*(5), 44–47.

Casillas A., Robbins S., Allen J., Kuo Yi., Hanson M., Schmeiser C. (2012). Predicting Early Academic Failure in High School from Prior Academic Achievement, Psychosocial Characteristics, and Behavior. *Journal of Educational Psychology*, 104(2), 407-420.

Flores M., Taylor M. (2007). The Effects of a Direct Instruction Program on the Fraction Performance of Middle School Students At-risk for Failure in Mathematics. *Journal of Instructional Psychology*, *34*(2), 84-94

Fredericks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59-109.

Georgakopoulos, I., Kytagias, C., Psaromiligkos, Y. Voudouri, A. (2018). Identifying risks factors of students' failure in e-learning systems: towards a warning system *International Journal of Decision Support Systems, 3*(3), 190-206.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate Data Analysis* (6th ed.). Upper Saddle River, New Jersey: Pearson Prentice Hall.

Hemmings B., Grootenboer P., Kay R. (2011). *International Journal of Science and Mathematics Education*, *9*(3), 691-705.

Hopf D., Xochellis P. (2003). *Gymnasium and Lyceum in Greece*. Athens. [in Greek].

Instructions on the Teaching of Mathematics at Gymnasium for the years 2018-2020. Ministry of Education. Available at: https://www.minedu.gov.gr/gymnasio-m-2/didaktea-yli-gymn/37395-28-09-18-odigies-gia-ti-didaskalia-ton-mathimatikon-ton-fysikon-epistimon-sto-gymnasio-gia-to-sxol-etos-2018-2020 [in Greek].

Kajander A., Zuke C., Walton G. (2008). Teaching Unheard Voices: Students At-Risk in Mathematics. *Canadian Journal of Education, 31*(4), 1039-1064.

Macfadyen L., Dawson S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computer and Education, 54*(2), 588-599.

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, *37,* 153-84.

McConney A., Perry L. (2010). Socioeconomic status, self-efficacy, and mathematics achievement in Australia: a secondary analysis. *Educational Research for Policy and Practice*, *9,* 77-91.

Sciarra, D.T. & Seirup, H.J. (2008). The multidimensionality of school engagement and math achievement among racial groups. *Professional School Counseling, 11*(4), 218-228. DOI: http://dx.doi.org/10.5330/PSC.n.2010-11.218

Smith, T. J., & McKenna, C. M. (2013). A comparison of the logistic regression pseudo Rsquared indices. *Multiple Linear Regression Viewpoints, 39*(2), 17-26.

Sullivan, P., Cheeseman, J., Michels, D., Mornane, A., Clarke, D., Roche, A. &Middleton, J. (2011). Challenging mathematics tasks: What they are and how touse them. In L. Bragg (Ed.), Maths is Multi-dimensional. *Proccedings of the 48thAnnual Conference of the Mathematical Association of Victoria*, 33–46. Melbourne: Mathematical Association of Victoria.

Tomlinson, C. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). Alexandria. VA: Association for Supervision and Curriculum Development.

Tomlinson, C. (2003). *Fulfilling the promise of the differentiated classroom: Tools and strategies for responsive teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Tzoka D., (2019). Mathematical challenge and diversified teaching in the math class. Diplomatic work. *Inter-university – interdepartmental postgraduate program "Didactics and Methodology of Mathematics"*.

Vose D. (2008). *Risk Analysis: A Quantitative Guide*, 3rd Edition. *Hoboken NJ Wiley*

Willms, J. D. (2003). *Student engagement at school: A sense of belonging and participation*. Paris: Organisation for Economic Co-Operation and Development.

Xin Y., Jiendra K., Buchman A., (2005). Effects of Mathematical Word Problem–Solving Instruction on Middle School Students with Learning Problems. *Journal of Special Education*, *39*(3),181-192.

Yurt E., (2014). The Predictive Power of Self-Efficacy Sources for Mathematics Achievement. *Education and Science*, *39*(176), 159-169