

Generating a Model to Predict Secondary School Students at Risk in Mathematics

Ioannis Georgakopoulos¹ , Stylianos Tsakirtzis^{1*} 

¹ University of West Attica, Department of Industrial and Product Design Engineering, GREECE

*Corresponding Author: s_tsakirtzis@hotmail.com

Citation: Georgakopoulos, I., & Tsakirtzis, S. (2021). Generating a Model to Predict Secondary School Students at Risk in Mathematics. *International Electronic Journal of Mathematics Education*, 16(2), em0630. <https://doi.org/10.29333/iejme/10877>

ARTICLE INFO

Received: 7 Jun. 2020

Accepted: 9 Jan. 2021

ABSTRACT

Mathematical courses aid individuals to deal with any problem that they might encounter in their life on the ground that Mathematics familiarize them with a problem-solving process. Additionally, people who have stood out in their profession are deemed to be those who have mastered their mathematical skills. 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. This paper proposes a way to predict secondary school students' critical performance in Mathematics through the generation of a potent model by means of a proper analysis of students' engagement data. A discriminant function analysis has been carried out on the respective data and the generated linear discriminant functions constitute the prediction model. The prediction model generation process is demonstrated through a case study on a specific Mathematical course delivered at a Greek private secondary school. The research outcome is very promising given that our model could potentially be used to develop an early warning system for secondary school students at risk in Mathematics.

Keywords: prediction model, students' engagement, students' final achievement, warning system

INTRODUCTION

The rudimentary knowledge on Mathematics is gained when a student is actively involved in the learning process (Carolan. & Guinn, 2007). 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.

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 parallel manner, the curriculum and syllabus list a number of factors and activities that should be considered in the process of secondary school students' evaluation in Mathematics. These factors and activities are cited 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 mastering and his critical thinking and creativity cultivation).
- The tasks performed by a student in the context of the daily learning process at school or at home, individually or collectively.
- Individual or group synthetic creative and interdisciplinary work.
- Hourly/few-minute written tests.
- Baccalaureate exams.

It is also essential to highlight that as dictated by the curriculum and syllabus the grade reflecting students' annual performance (graduation grade) could be computed as one third of the sum of the grades of the first quarter, the second semester and the written summary examination.

Our research interest is directed into secondary school students' behavioral engagement on the ground that this can be measured by graded activities in accordance to the curriculum and syllabus. Our research objective is to analyze the students' behavioral engagement data elicited from the respective activities defined in the light of the guidelines given on curriculum and

syllabus with a view to generating a model to predict secondary school students' critical performance. It is important to highlight that the term critical performance insinuates a threshold below which students at risk could be identified.

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 secondary 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 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. Another study (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.

Controlling the Risk of Students' Failure

In literature, we find many endeavors of controlling various risks, in many fields, through a warning system for those who are liable to face the underlying risk (Bouma & Van der Kaay, 1996; Huth et al., 2012; Kalkstein & Sheridan, 2007; Lambert et al., 2001; Lowe et al., 2010). In parallel manner, there are quite a few endeavors of controlling educational risks through a warning system that do we find in literature (Balfanz & Legters, 2004; Beck & Davidson, 2001; Macfayden & Dawson, 2010; Pinkus, 2008). In a more explanatory way, Balfanz and Legters (2004) have developed a warning system to provide high schools with the necessary feedback in regard to students who are at risk of dropping out, taking into consideration factors such as a poor course grade or a low-test score. Pinkus (2008) has come up with a warning system to improve graduation rates taking into account socioeconomic, individual, cultural and institutional factors. Beck and Davidson (2001) have developed an early warning system to predict low grades in college students from a specific survey of academic orientations, taking into consideration measures of intellectual skills or aptitudes and other non-intellectual factors such as undergraduates' perception of the academic environment. Though, there are many endeavors of predicting the learning performance of students which have not necessarily been led to a warning system development that do we find in literature. Ya-Han Hu et al. (2014) have provided a list with all these corresponding recent studies having them classified in terms of the type of data used, the attributes of such data and the techniques used to perform predictions.

Shedding more light on the method and the data used in various studies with a view to generating a prediction model for students at risk and developing an impending warning system we can stress on the effort of Macfayden and Dawson (2010), which is encircled on developing a warning system by applying statistical techniques to LMS tracking data captured at the end of the course. On the other hand, Hu et al. (2014) have developed a warning system by applying data mining techniques to LMS tracking data captured during the course. In a more elaborate detail, Macfayden and Dawson (2010) have used students' engagement LMS data, whereas Hu et al. (2014) have used learning portfolio LMS data. According to the study of Macfayden and Dawson (2010), the LMS tracking data which were proved to be statistically significant to constitute the base of a prediction model leading to a warning system were non-time related course material data. On the contrary, according to the study of Hu et al. (2014), the LMS corresponding tracking data which were proved to be statistically significant were time-related course material data. It is vital to highlight that the prediction models developed and the impending warning systems in all referred studies are dependent on the course structure and achieve the intended result in terms of a specific course.

Summing up, there are not many endeavors of developing a warning system for students at risk in Mathematics that do we find in literature. In parallel manner, there are not many endeavors of predicting secondary school students at risk that do we meet in literature. Thereby, there is great scope for scientific output. In this paper, we demonstrate a prediction model for secondary school students at risk in mathematics by carrying out a discriminant function analysis on students' behavioral engagement data (Georgakopoulos et al., 2018). The prediction model is demonstrated through a case study on a specific Mathematical course delivered at a Private Secondary School. Needless to say, the prediction model verification is in the pipeline. The next section unveils our method and the pillar on which our prediction model has been generated.

Table 1. Data Collected

Average Theory Test Score
Percentage of Exercises Completed
Average Homework Score
Percentage of Review Packages Completed
Average Review Packages Score
Percentage of School Work Completed
A' Semester Rating
B' Semester Rating
Absences
Score of Final Examination
Graduation Grade

Table 2. Data Measurement Time

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
Absences	Daily

METHOD

According to Vose (2008), the prediction model is part of a risk analysis process which is also part of a generic risk management framework. In the spirit of such framework and in the light of the methodology demonstrated in the work (Georgakopoulos et al., 2018) we propose a concrete methodology which includes three stages:

1. Data Collection
2. Prediction Model Generation
3. Prediction Model Verification

The data collection process includes the gathering of all data in regard to students' behavioral engagement elicited in the context of any activity that could reflect the overall students' involvement in the learning process as dictated by the curriculum and syllabus. These data are afterwards analyzed in terms of a discriminant function analysis (Georgakopoulos et al., 2018) in order to come up with a prediction model for students' critical performance. That prediction model includes two classification functions, one for students at risk and one for students who pass. The score of both function is calculated for each student before the final exams and the maximum score determines the classification group of each student.

The prediction model should be verified in terms of its accuracy (correct classification percentage) being applied to a plethora of courses sharing the same learning design and the verified prediction model could set the standards for an early warning system for secondary school students at risk in Mathematics. The prediction model generation is demonstrated in the context of this paper but the verification process is not demonstrated on the ground that it is in progress.

In a more elaborate detail, in the spirit of the curriculum and syllabus, we gathered the respective collective data out of three grades (Grade A; Grade B and Grade C) in respect to a mathematical course delivered at a private Secondary School. Needless to say, 152 students were enrolled into the course. The data set collected is listed into **Table 1: Table 2** indicates when the data set was measured.

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). All the variables which were modeled to describe the respective data are listed into **Table 3**. The variables modeled were deployed in terms of a discriminant function analysis in order to come up with a prediction model of students' at risk (Georgakopoulos et al., 2018) based on the classification functions which have already been referred.

Table 3. Variables Modeled

Data Description	Variable Modeled
Average Theory Test Score	TTS
Percentage of Exercises Completed	PEC
Average Homework Score	HS
Percentage of Review Packages Completed	PRPC
Average Review Packages Score	RPS
Percentage of School Work Completed	PSWC
A' Semester Rating	ASR
B' Semester Rating	BSR
Absences	Ab
Score of Final Examination	SFE
Graduation Grade	GG
Students at risk	srisk

Table 4. Linear Discriminant Functions' Coefficients

	Classification Function Coefficients	
	0	1
PEC	11,410	-2,732
PRPC	-3,867	-3,433
PSWC	-62,009	-61,590
RPS	,112	,123
Ab	,549	,703
TTS	,771	,656
HS	4,607	4,861
ASR	-,727	-,860
BSR	1,773	2,089
(Constant)	-32,456	-34,025

Table 5. Classification Table

Classification Results ^a					
		srisk	Predicted Group Membership		Total
			0	1	
Original	Count	0	134	2	136
		1	6	10	16
	%	0	98,5	1,5	100,0
		1	37,5	62,5	100,0

RESULTS

Table 4 indicates the coefficients of the generated linear discriminant functions for students at risk and for students who are about to pass the course.

The field "0" on the column "srisk" provides us with the linear discriminant functions' coefficients for students who are about to pass the course whereas the field "1" on the column "srisk" provides us with the corresponding coefficients for students who are about to fail the course. The correct classification percentage achieved by the linear discriminant functions is clarified in **Table 5**. According to the respective table the discriminant functions achieve a 94.7 % correct classification percentage.

Figure 1 graphically illustrates the classification outcome for students who are about to pass the course whereas **Figure 2** graphically depicts the classification outcome for students at risk.

Figure 1 and **Figure 2** show that the discriminant functions work better on the case of students at risk.

CONCLUSIONS AND FUTURE EXPECTATIONS

We have come up with a potent prediction model which achieves a high classification percentage. The only defect in our model is that the generated linear discriminant functions work better on the case of students at risk and thereby our model could potentially be ameliorated.

Our model has been developed in terms of a specific mathematical course and its accuracy has been tested in the context of the respective course. That finding is in line with the studies referred in literature where the generated prediction models were tailored to the needs of a specific course. However, our model could potentially be applied to a plethora of courses having the same learning design in order to come up with a prediction model suitable for many similar courses.

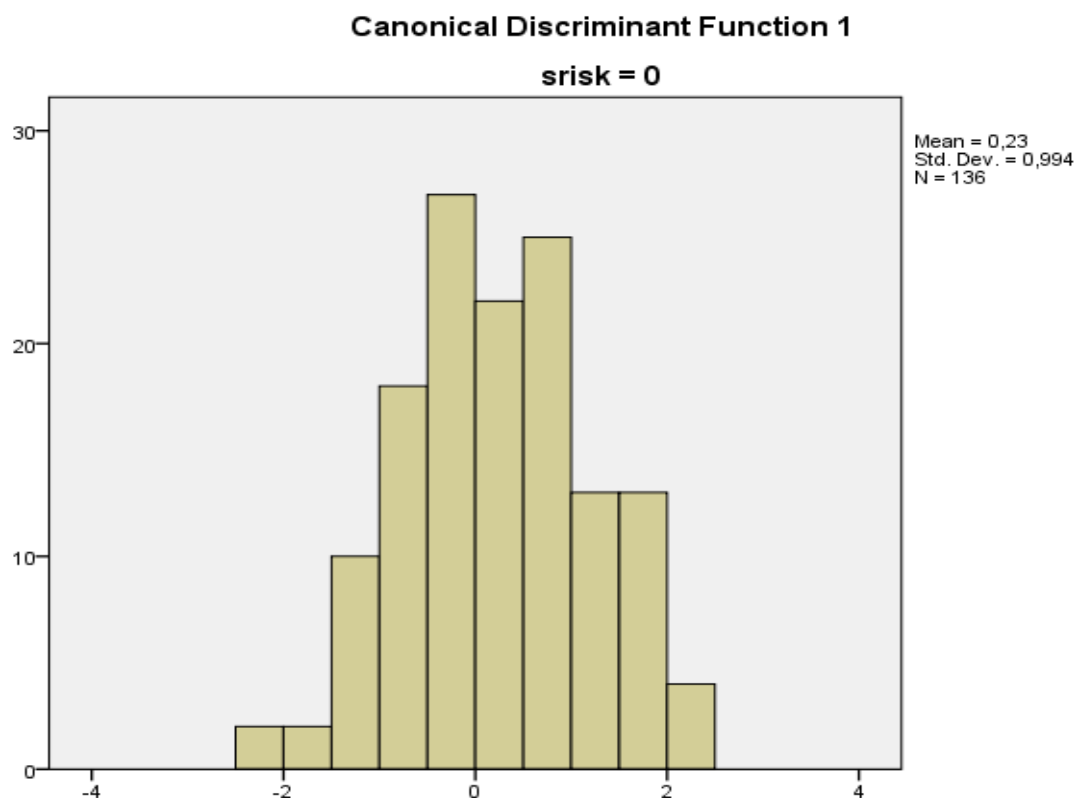


Figure 1. Graphical Illustration for the classification outcome (Students who pass the course)

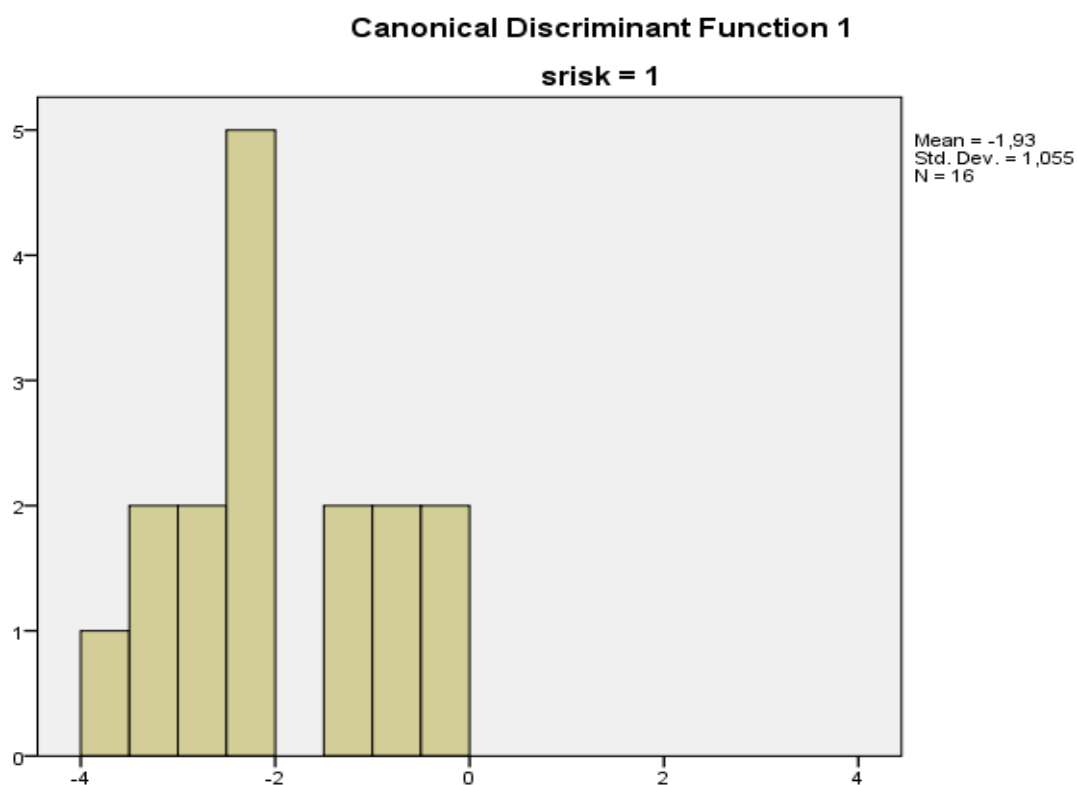


Figure 2. Graphical Illustration for the classification outcome (Students who fail the course)

Our research does not draw attention into social, psychological or economical factors on the ground that the data analyzed are correlated to the students' behavioral engagement as dictated by the curriculum and syllabus. Nevertheless, our model could be additionally applied to a plethora of courses in the light of the previously cited factors to assess the impact of those factors on the students' achievement.

It is also essential to highlight that currently we are working on verifying the prediction model in order to develop a warning system that could be used by teachers to control the risk of students' failure in regard to the underlying mathematical course. In parallel manner, the prediction model could be tested in the context of a plethora of similar courses with a view to developing an early warning system suitable for all courses sharing the same learning design.

Author contributions: All authors have sufficiently contributed to the study, and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Declaration of interest: No conflict of interest is declared by authors.

REFERENCES

- Balfanz, R., & Legters, N. (2004). Locating the dropout crisis: Which high schools produce the nation's dropouts? In G. Orfield (Ed.), *Dropouts in America* (pp. 57-84). Harvard Education Press.
- Beck, H., & Davidson, W. (2001). Establishing an Early Warning System: Predicting Low Grades in College Students from Survey of Academic Orientations Scores. *Research in Higher Education*, 42(6), 709-723. <https://doi.org/10.1023/A:1012253527960>
- Bouma, M., & Van der Kaay, H. (1996). The El Nino Southern Oscillation and the historic malaria epidemics on the Indian subcontinent and Sri Lanka: an early warning system for future epidemics? *Tropical Medicine and International Health*, 1(1), 86-96. <https://doi.org/10.1046/j.1365-3156.1996.d01-7.x>
- Carolan, J., & Guinn, A. (2007). Differentiation. *Lessons from Master Teachers. Educational Leadership*, 64(5), 44-47.
- Casillas, A., Robbins, S., Allen, J., Kuo, Y., 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. <https://doi.org/10.1037/a0027180>
- 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. <https://doi.org/10.3102/00346543074001059>
- Georgakopoulos, I., Kytasias, 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(3/4), 190. <https://doi.org/10.1504/IJDSS.2018.100187>
- Hemmings, B., Grootenboer, P., & Kay, R. (2011). Predicting mathematics achievement: The influence of prior achievement and attitudes. *International Journal of Science and Mathematics Education*, 9(3), 691-705. <https://doi.org/10.1007/s10763-010-9224-5>
- Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, 28(5), 524-549. <https://doi.org/10.2307/749690>
- Hopf, D., & Xochellis, P. (2003). *Gymnasium and Lyceum in Greece*. Athens. [in Greek].
- Hu, Y.-H., Lo, C.-L., & Shih, S.-P. (2014). Developing early warning systems to predict students' online learning performance. *Journal: Computers in Human Behavior*, 36(C), 469-478. <https://doi.org/10.1016/j.chb.2014.04.002>
- Huth, V., Biral, F., Martin, O., & Lot, R. (2012). Comparison of two warning concepts of an intelligent Curve Warning system for motorcyclists in a simulator study. *Accident Analysis and Prevention*, 44, 118-125. <https://doi.org/10.1016/j.aap.2011.04.023>
- Kajander, A., Zuke, C., & Walton, G. (2008). Teaching Unheard Voices: Students At-Risk in Mathematics. *Canadian Journal of Education*, 31(4), 1039-1064.
- Kalkstein, A., & Sheridan, C. (2007). The social impacts of the heat-health watch/warning system in Phoenix, Arizona: assessing the perceived risk and response of the public. *International Journal of Biometeorology*, 52, 43-55. <https://doi.org/10.1007/s00484-006-0073-4>
- Lambert, J., Finch, A., & Schaalje, B. (2001). Psychotherapy Quality Control: The Statistical Generation of Expected Recovery Curves for Integration in an Early Warning System. *Clinical Psychology and Psychotherapy*, 8, 231-242. <https://doi.org/10.1002/cpp.286>
- Lowe, R., Bailey, T., Stephenson, D., Graham, R., Coelho, C., Carvalho, M., & Barcellos, C. (2010). Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in Brazil. *Computers & Geosciences*, 37(3), 371-381. <https://doi.org/10.1016/j.cageo.2010.01.008>
- 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. <https://doi.org/10.1016/j.compedu.2009.09.008>
- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, 37, 153-184. <https://doi.org/10.3102/00028312037001153>
- 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. <https://doi.org/10.1007/s10671-010-9083-4>
- Pinkus, L. (2008). Using early-warning data to improve graduation rates: Closing cracks in the education system. *Alliance for Excellent Education. Policy Brief*. <https://all4ed.org/wp-content/uploads/EWI.pdf>

- Sciarra, D., & Seirup, H. (2018). The multidimensionality of school engagement and math achievement among racial groups. *Sagepub Journal*, 11(4), 218-228. <https://doi.org/10.5330/PSC.n.2010-11.218>
- 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. Proceedings of the 48th Annual Conference of the Mathematical Association of Victoria* (pp. 33-46). Mathematical Association of Victoria.
- Tomlinson, C. (2001). *How to differentiate instruction in mixed-ability classrooms* (2nd ed.). Association for Supervision and Curriculum Development.
- Tomlinson, C. (2003). *Fulfilling the promise of the differentiated classroom: Tools and strategies for responsive teaching*. 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 Ed.). Wiley.
- Willms, J. D. (2003). *Student engagement at school: A sense of belonging and participation*. 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. <https://doi.org/10.1177/00224669050390030501>
- Yurt, E. (2014). The predictive power of self-efficacy sources for mathematics achievement. *Education and Science*, 39(176), 159-169. <https://doi.org/10.15390/EB.2014.3443>