DETERMINANTS IN MODELLING EARLY SCHOOL DROPOUT

Bianca-Raluca Cibu [1*], Camelia Delcea [2], Adrian Domenteanu [3]

[1]Bucharest University of Economic Studies, Bucharest, Romania, cibubianca18@stud.ase.ro
[2]Bucharest University of Economic Studies, Bucharest, Romania, camelia.delcea@csie.ase.ro
[3]Bucharest University of Economic Studies, Bucharest, Romania, domenteanuadrian18@stud.ase.ro

Abstract
Recognizing the importance of continuity in education, it was deemed necessary to carry out a study whose main objective is to identify the factors that lead to the intention to drop out in order to formulate effective strategies to combat this phenomenon. Early school dropout has negative consequences both for individual development and for social and economic progress. Therefore, this paper aims to contribute to the understanding and prevention of this phenomenon through a rigorous analysis of its determinants. In this context, eight relevant dependent variables have been identified in the literature that are believed to play a significant role in the intention to drop out of school. These variables include factors such as school absenteeism, alcohol or substance abuse, attitude, awareness, family, family supervision, school environment and school rules. The analysis used in the study examines these significant variables through structural equation modeling (SEM). SmartPLS software was used to conduct this analysis, which allows the use of Partial Least Squares SEM (PLS-SEM) and Bootstrapping modeling techniques. The data used for this research was collected using a well-structured questionnaire consisting of 28 questions aimed at capturing students' perceptions and experiences of school and the factors that might contribute to their intention to drop out of school. A total of 669 respondents completed the questionnaire, providing a solid database for analysis.

Keywords: PLS-SEM, Bootstrapping, Romania, education, EU periphery

JEL Classification: H75, I21, I25, P36

DOI: 10.24818/CTS/6/2024/2.02

1. Introduction

Analyzing the intention to drop out of school is essential as it provides the opportunity to intervene before students actually drop out, thus reducing the long-term risks associated with dropping out. It enables the identification of underlying factors such as academic difficulties, lack of motivation, emotional problems or a disadvantaged socio-economic background. A better understanding of these factors enables the development of early interventions tailored to the needs of each individual student, such as psychological counseling, additional academic support or extracurricular activities aimed at improving students' motivation and sense of belonging at school. Preventing early school leaving also helps to reduce social exclusion and increases the chances of integration into the labor market, which has a positive impact on both the individual and society as a whole.

Adequate education reduces the burden of disease, increases life expectancy, delays ageing, reduces health inequalities and reduces unhealthy habits (Freudenberg and Ruglis, 2007). Over time, it has been shown that people with higher levels of education are more likely to be in better health (Winkleby et al., 1992). Lower education predicts earlier death, while higher formal education is consistently associated with lower mortality rates (Molla, Madans and Wagener, 2004).

In the context of education, the term dropout refers to the phenomenon of students leaving school before completing their studies. The dropout rate is a very important indicator, expressed in money or as a percentage of students who leave the education system early. Various studies have been conducted

* Corresponding author
Authors' ORCID:
Cibu Bianca-Raluca
Delcea Camelia
Domenteanu Adrian

showing the prevalence of dropout at different levels of education, including primary and secondary (Haimovich, Vazquez and Adelman, 2021).

In the field of online education, the complexity of defining dropout is even greater, as certain factors such as family or work constraints can influence a student's decision to continue their studies or not (Grau-Valldosera and Minguillón, 2014).

The high dropout rate among students is a major problem that severely affects education systems worldwide. It has a negative impact not only on educational institutions, which lose valuable resources and face a decline in overall academic performance, but also on society as a whole. The impact is reflected in a less educated, lower-skilled workforce, which reduces economic competitiveness and increases pressure on social systems. Furthermore, the impact on the individual student is profound, as dropping out of school negatively affects both personal development and career opportunities, limiting young people's ability to improve their life chances (Hassan, Muse and Nadarajah, 2024).

In this paper we have started with the presentation of the literature in order to identify the variables that we will use in the construction of the questionnaire, and later in the methodology we will explain all the indicators that we will take into account in the construction of the model. The next step is the actual analysis, in which Partial Least Squares SEM (PLS-SEM) and Bootstrapping modeling techniques were applied to a database created using a questionnaire with the help of the SmartPLS program. In the final phase, some conclusions are drawn based on the results obtained. The main objective of the analysis was to identify the factors that influence school dropout.

2. Literature review

Chronologically, school dropout could be seen as the last phase of a dynamic process in which several causes at family, school or individual level could describe the phenomenon (Bradshaw, O'Brennan and McNeely, 2008; De Witte et al., 2013).

School absenteeism is associated with a number of adverse conditions, such as teenage pregnancy, psychiatric disorders, inappropriate behavior, delinquency, and abuse of alcohol, tobacco, marijuana, or other substances (Nik Jaafar et al., 2013). In addition, young people who are excessively truant are at high risk of dropping out of school permanently, which can lead to economic deprivation and various psychological, social, occupational and marital problems in adulthood (Kogan et al., 2005). Numerous research studies have been conducted to identify risk factors for school absenteeism and school dropout. Some of these risk factors are related to child characteristics (child's age, internalizing problems, externalizing problems and poor physical health), parental characteristics (parental mental health problems and parental unemployment), family characteristics (low socioeconomic status and family breakdown) or school characteristics (poor quality of teacher training) (Gubbels, van der Put and Assink, 2019).

Patterson et al. (1991), cited by Gaik et al. (2010) found that children with early behavioral problems are often at increased risk of developing academic difficulties and being rejected by high-performing peers who could provide a positive role model. This exclusion can lead them to fall into groups of behaviorally challenged peers, which in turn leads to behaviors such as truancy, drug use or violence. On the other hand, students who abide by school rules tend to perform better academically and are less likely to drop out of school (Bradshaw, O'Brennan and McNeely, 2008). In addition, disruptive behavior is influenced by the level of parental involvement and the relationships between teachers and students, which can increase the impact on academic performance.

Substance abuse is a significant individual risk factor for dropping out of school. The relationship between drug or alcohol use and school dropout is one of the most frequently studied in the literature (Esch et al., 2014), suggesting that students involved in substance abuse are more likely to drop out of school (Bradshaw, O'Brennan and McNeely, 2008). Esch et al. (2014) found that students who continue their education have a significantly lower risk of becoming addicted to alcohol than their peers who drop out of school.

Parental divorce and family conflict can affect student behavior both inside and outside the classroom (Bradshaw, O'Brennan and McNeely, 2008). Family structure plays an important role in the predisposition to dropping out of school, with studies showing that students from single-parent families

are at higher risk of dropping out (Bridgeland, DiIulio and Morison, 2006; De Witte et al., 2013). A lack of rules and parental neglect, especially in families with a low level of education, can impair the socialization process and lead to risky behavior such as the consumption of alcohol or other substances (Park and Kim, 2016).

The longer the time spent on the way to school, the less attractive it can be for the child (Afoakwah and Koomson, 2021; Ding and Feng, 2022). Accompanying the child to school can be a major problem for parents (Lidbe et al., 2020). Long travel time affects the student's participation in class and can lead to dropping out of school, especially for low-income students (Ramírez-Hassan et al., 2023).

Using data from the 2022 National Education Affordability Survey (NEAS), Hassan et al. (2024) suggested that machine learning techniques predict dropout rates. In their analysis, they found that age, gender and grades achieved were among the variables that had a significant impact. Hancco-Monrroy et al (2024), aware of the time of the COVID-19 pandemic, investigated the relationship between professionalism in medical schools and the intention to drop out. Factors identified as having a positive impact on dropping out were working in the private sector, depression, anxiety and the ability to work in a team. In contrast, factors that were found to have a negative correlation with dropping out were studying at a public institution, the ability to learn and subjective well-being.

In their study, Berens et al. (2018) focused on comparing two data sets from universities, one from the private sector and one from the public sector. They concluded that the further students progress in their studies, the more the level of academic achievement is a robust indicator of dropout. Theodorou et al. (2024) followed a pilot study whose main objective was to evaluate a program designed to prevent the risk of mental health problems, an intervention that was evaluated using a questionnaire completed by participants before and after the study. As for the results, an improvement in emotion regulation and a reduction in externalizing and internalizing problems were observed. The study concluded that this type of intervention can reduce school dropout.

3. Methodology

In this part of the realized work, the methodology of data collection and the analysis techniques used to observe the results are presented.

For the data collection used in the analysis, a questionnaire with a total of 28 questions (Appendix A: Research Questionnaire) was created based on variables:

- *Absenteeism from school (ABS),* which is considered the first step towards dropping out of school, has serious consequences such as poor performance, social maladjustment and problems in adulthood. Recovery is difficult as the student loses contact with peers and the subject matter;
- *Alcohol or substance abuse (AB)* is attracting more and more young people and affecting their school attendance, concentration and academic performance;
- *Attitudes (AT)* towards dropping out of school reflect people's perception of the problem: some see it as a social problem, others are indifferent or consider education and school attendance unimportant. It is also about involvement in preventing dropout by supporting those who tend to drop out of school;
- Students need to be *aware* of the consequences of dropping out of school *(AW),* as knowledge of the risks and difficulties can influence the decision to drop out. This variable measures how aware they are of the negative consequences and how they perceive a life without education;
- *Family (FAM),* social class and unpleasant events (quarrels, divorce) can influence students' academic performance. The family plays an essential role, as the parents' ideas and values regarding education are passed on to the children;
- *Family monitoring (FM):* Some students will only learn if they are motivated by their parents or guardians, and neglect of school must be carefully monitored. The family plays a crucial role in a child's education and a lack of involvement can lead to dropping out of school, especially for students who need constant support;

- The *school environment (SE),* which integrates students into a close-knit group, reduces the risk of dropping out of school, even in difficult situations. However, conflicts or a lack of adaptation to the collective can lead to a refusal to attend school;
- *School rules (SR),* a demotivating factor for students is the discrepancy between the textbooks and the subject matter, as well as the desire to learn current affairs. Although ten classes are compulsory, the lack of penalties for dropping out of school makes the process easy. Stricter school regulations could encourage the completion of compulsory classes;

The intention to drop out of school (INT) is the last step before making a clear decision, which is influenced by internal and external factors. The questions can be used to analyze students' plans for completing their studies and any doubts they may have. The variables used in the questionnaire were identified in the literature, as shown in Table 1. Each of these variables was associated with 3 or 4 questions whose possible answers were rated on a Likert scale: 1 - To a very low extent, 2 - To a low extent, 3 - Neutral, 4 - To a high extent, 5 - To a very high extent. The survey was conducted using Google Forms and distributed to students from different groups in order to collect as many relevant responses as possible and thus obtain a wider age range.

Table 1. Variables used and selected references

	ABS	AB	AT	AW	FAM	FM	INT	SE	SR
De Witte et al. (2013)			x	x	x	x	x	x	x
Fernández-Suárez et al. (2016)	x	x				x		x	
Bradshaw et al. (2008)		x	x	x	x	x	x	x	x
Esch et al. (2014)		x	x						
Park & Kim (2016)	x	x		x	x	x		x	x
Bridgeland et al. (2006)					x	x		x	x

Source: Authors processing

Once the data had been collected, it was entered into the SmartPLS program, which is used for partial structural equation modeling (SEM). This modeling, a multivariate data analysis technique, is often used to test additive and linear causal models (Chin, Marcolin and Newsted, 2003; Haenlein and Kaplan, 2004). Within a SEM, there are two sub-models: the inner model for the relationships between independent and dependent latent variables and the outer model for the relationships between latent variables and observed indicators. PLS-SEM has the ability to combine both structural and measurement models to assess the complex relationships between variables. On the other hand, PLS-SEM is a combination of approximate regression analysis and minimum hypothesis factor analysis; the R-squared index obtained from the model represents the extent to which the independent variable explains the dependent variable (Hair, Ringle and Sarstedt, 2011).

To ensure the validity of the model, convergent validity had to be examined first (Hair Jr. et al., 2014). To ensure this, the external loading of each indicator must be greater than 0.70 and the average variance extracted (AVE) of each variable must be equal to or greater than 0.50. The AVE indicates how high the mean squared loading of a group of indicators is. If the AVE is equal to 0.50, this means that the construct explains at least 50% of the variance of its indicators (Hair et al., 2014).

Internal consistency or composite reliability is a measure of the internal consistency between the scale items (Netemeyer, Bearden and Sharma, 2003). The higher the value of this indicator, the higher the reliability (Jöreskog, 1971). If the values are between: 0.60-0.70 we have an acceptable reliability, 0.70-0.90 satisfactory to good, and values above 0.95 may prove problematic, with low construct validity (Diamantopoulos et al., 2012). Reliability can also be measured using Cronbach's Alpha, but is not as precise and requires similar thresholds, but has lower values compared to composite reliability. In addition, there is also the "rho_A" index, which reflects the reliability of constructs almost exactly and often lies between the composite reliability and Cronbach's Alpha as value (Dijkstra and Henseler, 2015).

The variance inflation factor (VIF) is also analyzed to determine statistical collinearity. Ideally, the value for this indicator should be around 3 or even lower (Mason and Perreault, 1991). R-squared indicates the proportion of the variance of a dependent variable that is explained by an independent variable. R-squared values range from 0 to 1 and are usually expressed as a percentage. In the case of

multiple regression, where there are many independent variables, the R-squared must be adjusted. The adjusted R-squared compares the descriptive power of regression models that contain a different number of predictors. Each predictor that is added to a model increases the R-squared and does not constantly decrease it. It only increases if the new term improves the model (Fernando, 2024).

After reviewing and eliminating values that do not comply with the above rules, we must ensure that at least two indicators remain for each variable that represent it (Cheah et al., 2018), although the cases in which individual items are accepted are special cases. In these exceptional cases, the mean variance or compositional reliability cannot be analyzed (Hair et al., 2014).

The use of bootstrapping has the advantage that the statistical hypotheses are not restrictive, which is an advantage for the user, since in most cases the empirical data do not fulfill the restrictive hypotheses (Mooney and Duval, 1993). This technique was used to analyze the statistical significance of the analysis performed with the PLS algorithm. The variables are statistically significant if the P-value is less than 0.05 at a 95% confidence interval.

4. Results

In order to understand what factors lead young people to drop out of school, a number of variables were identified with the help of specialist literature, which formed the basis for the creation of a questionnaire with 28 questions (Appendix A: Research questionnaire), which was subsequently completed by a total of 669 respondents. Some details of the participants can be found in Table 2. As far as the age groups are concerned, the 12 to 25-year-olds predominate with a total of 603 respondents, who represent our target group to be analyzed. As expected, the number of women who took part in this study is significantly higher than that of men.

Table 2. Participants details

Socio-economic variables	Variants	Frequency	Percentage
Age	6-11 Years	24	4%
	12-15 Years	66	10%
	16-19 Years	192	29%
	20-25 Years	345	52%
	>25 Years	42	6%
Educational level	Primary education	23	3%
	Secondary education	191	29%
	Post-secondary education	4	1%
	Undergraduate studies	376	56%
	Master studies	36	5%
	Doctoral studies	16	2%
Gender	Female	465	70%
	Male	204	30%

Source: Authors processing

In order to carry out the analysis, the data was entered into the Smart PLS software. The variables identified and used in the analysis are: absenteeism, alcohol or substance abuse, attitude, awareness, family, family supervision, intention to drop out, school environment and school rules. A 5-point Likert scale was used to control the responses: 1 - To a very low extent, 2 - To a low extent, 3 - Neutral, 4 - To a high extent, 5 - To a very high extent. Both the variables and the number of questions for each can be seen in Figure 1, where the program was used to illustrate the initial model of analysis.

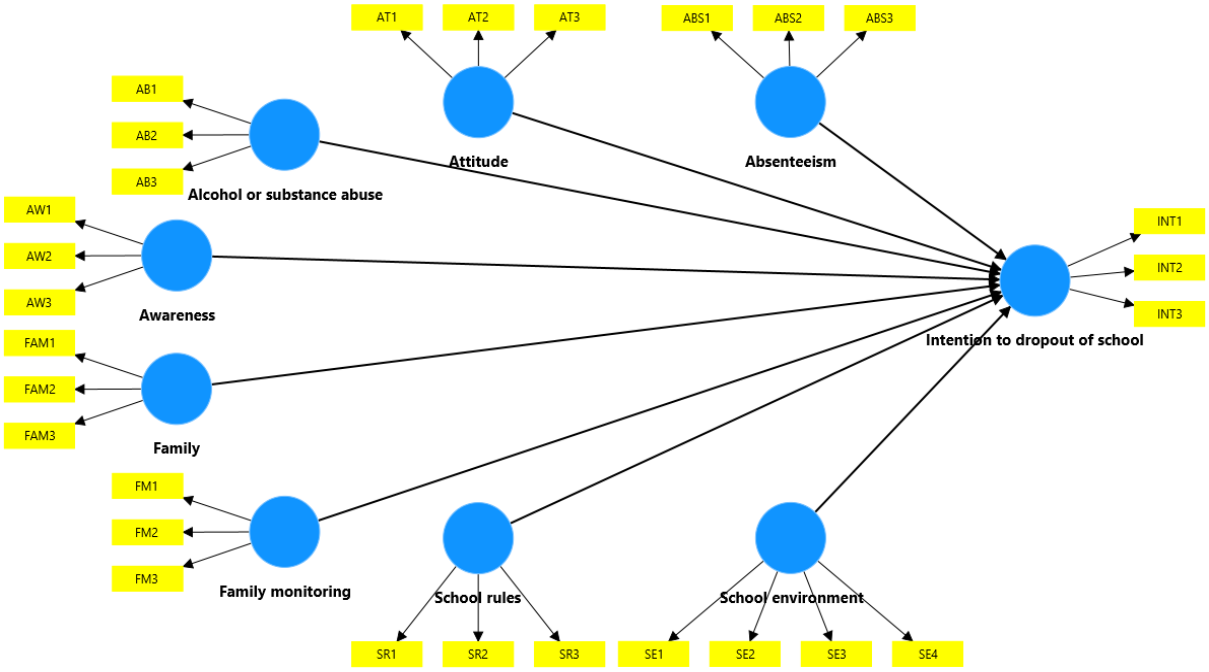

Fig. no. 1. Initial model
Source: Authors processing

After creating the initial model, some indicators are analyzed according to the methodology presented in the previous section to ensure that our model is valid. For the initial model, the values for the outer loadings are presented in Table 3. According to the researchers, the accepted value for the outer loadings is at least 0.70, which means that the related indicators have a lot in common and explain more than 50% of the variance of the indicator, but the table also shows lower values. These values are gradually removed from the model, starting with the lowest value. The highest loading of 0.999 is found between the indicator SR3 and the variable school rules.

Table 3. Outer loadings - initial model

	ABS	AB	AT	AW	FAM	FM	INT	SE	SR
AB1		0.877							
AB2		0.909							
AB3		0.935							
ABS1	0.940								
ABS2	*-0.102*								
ABS3	*0.404*								
AT1			0.864						
AT2			0.883						
AT3			0.835						
AW1				0.895					
AW2				0.882					
AW3				0.887					
FAM1					*0.633*				
FAM2					0.928				
FAM3					*0.681*				
FM1						*0.617*			
FM2						*-0.189*			
FM3						*0.019*			
INT1							0.910		
INT2							0.912		
INT3							0.943		
SE1								*-0.118*	
SE2								*0.317*	
SE3								*0.022*	

	ABS	AB	AT	AW	FAM	FM	INT	SE	SR
SE4								0.959	
SR1									*0.426*
SR2									*0.592*
SR3									0.999

Source: Authors processing

In parallel, AVE, Cronbach's Alpha, rho_A and composite reliability are analyzed (Table 4). Taking into account the specifications of the methodology, the absenteeism, variable is not validated by any indicator. The composite reliability is below 0.6, the AVE is below 0.5, and the two additional indicators of reliability, Cronbach's Alpha and rho_A, are below 0.7. Each variable is checked in this way and outliers are deleted.

Table 4. Building reliability and validity - initial model

	Cronbach's Alpha	Composite Reliability (rho_A)	Composite Reliability	Average Variance Extracted (AVE)
Absenteeism	*0.501*	*-0.008*	*0.442*	*0.352*
Alcohol or substance abuse	0.893	0.906	0.933	0.823
Attitude	0.826	0.832	0.896	0.741
Awareness	0.866	0.866	0.918	0.788
Family	0.709	1.154	0.798	0.575
Family monitoring	0.774	*-3.728*	*0.072*	*0.139*
Intention to dropout of school	0.911	0.912	0.944	0.850
School environment	*0.660*	*0.007*	*0.320*	*0.259*
School rules	0.720	13.902	0.735	0.510

Source: Authors processing

As for the VIF, its value should be 3 or even lower, which is not the case for AB3 and INT3 (Table 5). This indicator helps us to analyze the degree of multicollinearity. It is present when there is a correlation between several independent variables in a multiple regression model.

Table 5. Collinearity statistics (VIF) - initial model

	VIF
AB1	2.452
AB2	2.621
AB3	*3.327*
ABS1	1.068
ABS2	1.176
ABS3	1.241
AT1	1.792
AT2	2.108
AT3	1.823
AW1	2.314
AW2	2.195
AW3	2.225
FAM1	1.511
FAM2	1.246
FAM3	1.613
FM1	1.835
FM2	1.871
FM3	1.371
INT1	2.958
INT2	2.908
INT3	*3.955*
SE1	1.653

	VIF
SE2	1.473
SE3	1.652
SE4	1.060
SR1	1.279
SR2	1.565
SR3	1.533

Source: Authors processing

By excluding all indicators whose values did not fall within the ranges specified in the methodology, the following final model was obtained from Figure 2. In contrast to the initial model, this time we can observe that our dependent variable, intention to drop out of school, is only influenced by the variables attitude, alcohol or substance abuse and awareness, with the number of indicators for each of these variables ranging from 2 to 3 (questions).

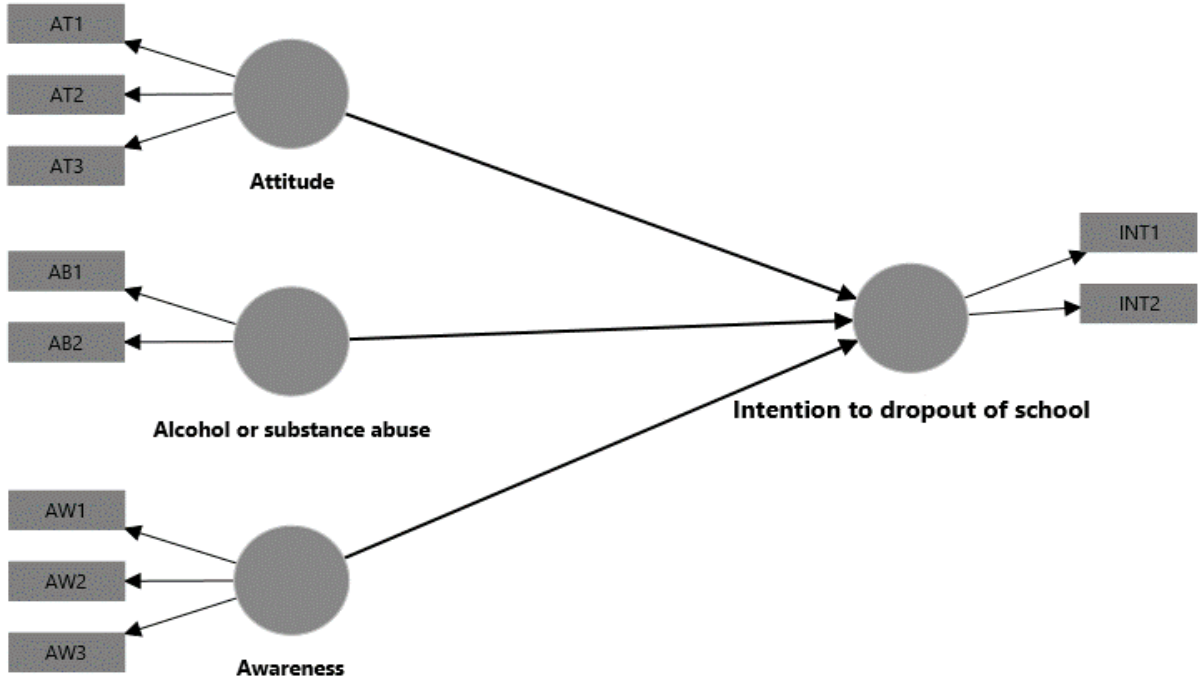

Fig. no. 2. Final model
Source: Authors processing

To ensure the validity of the model, the loading is presented in Table 6, AVE, Cronbach's Alpha and Composite Reliability in Table 7, and the VIF values for each question in Table 8.

Table 6. Outer loadings - final model

	Alcohol or substance abuse	Attitude	Awareness	Intention to dropout of school
AB1	0.898			
AB2	0.931			
AT1		0.863		
AT2		0.882		
AT3		0.837		
AW1			0.895	
AW2			0.882	
AW3			0.886	
INT1				0.925
INT2				0.930

Source: Authors processing

Cronbach's alpha values ranging from 0.806 to 0.866 indicate a fairly high internal consistency. Composite reliability values of over 0.831 also confirm the reliability of the constructs. All AVE values

above 0.741 indicate strong convergent validity, which emphasizes that the indicators explain a significant part of the variance of the variables. The metrics used ensure that the constructs are analyzed with consistency and accuracy.

Table 7. Building reliability and validity - final model

	Cronbach's Alpha	Composite Reliability rho_A	Composite Reliability	Average Variance Extracted (AVE)
Alcohol or substance abuse	0.806	0.825	0.911	0.837
Attitude	0.826	0.831	0.896	0.741
Awareness	0.866	0.866	0.918	0.788
Intention to dropout of school	0.837	0.838	0.925	0.860

Source: Authors processing

The VIF values between 1.792 and 2.314 indicate that the variables of the model are moderately correlated.

Table 8. Collinearity statistics (VIF) - final model

	VIF
AB1	1.840
AB2	1.840
AT1	1.792
AT2	2.108
AT3	1.823
AW1	2.314
AW2	2.195
AW3	2.225
INT1	2.074
INT2	2.074

Source: Authors processing

Both the R-squared and the adjusted R-squared are quite high, both at 79%. Thus, 79% of the variation in dropout intention can be explained by the variables used in the model.

Table 9. R-square - final model

	R-square	R-square adjusted
Intention to dropout of school	0.786	0.785

Source: Authors processing

The Bootstrapping algorithm was then used to calculate the index of the p-values in order to determine whether the existing relationships between the variables are statistically significant. The following assumptions were made:

- **H1**: *The abuse of alcohol/prohibited substances* influences the intention of pupils/students to drop out of school.
- **H2**: *Pupils'/students' attitudes* towards the problem under investigation and towards education influence their intention to drop out of school.
- **H3**: Young people's *awareness* of the importance and role of education for their future affects their intention to drop out of school.

Analyzing the significance from a statistical point of view, we can see in Table 10 that only H2 and H3 are significant, while H1 is not significant because the coefficient for the p-value is greater than 0.05. For all three hypotheses, we observe negative values for the path coefficients, which means that the intention to drop out of school decreases as one of the three variables increases.

Table 10. Validating hypotheses

Hypotheses	Relationship	Path Coefficients	Structural path coefficients	Decision
H1	Alcohol or substance abuse -> Intention to dropout of school	-0.046	*0.062*	Not Supported
H2	Attitude -> Intention to dropout of school	-0.185	0.000***	Supported
H3	Awareness -> Intention to dropout of school	-0.690	0.000***	Supported

Source: Authors processing
*Note: ***p ≤ 0.001*

Following all the previous steps and arriving at the same significant variables, Table 11 shows a breakdown of the analysis by gender. This time we find that, from a statistical point of view, all three hypotheses are statistically significant for the male respondents, in contrast to the female respondents, for whom, as in the initial model, only H2 and H3 are significant.

Table 11. Validating hypotheses based on gender

Hypotheses	Relationship	Structural path coefficients		Significance
		Female	Male	
H1	Alcohol or substance abuse -> Intention to dropout of school	0.829	0.008*	Male
H2	Attitude -> Intention to dropout of school	0.017*	0.022*	Both
H3	Awareness -> Intention to dropout of school	0.000***	0.000***	Both

Source: Authors processing
*Note: *p ≤ 0.05, ***p ≤ 0.001*

It is very interesting to note that both at the gender-specific level and at the level of the overall analysis, the greatest influence on the intention to drop out of college is the degree of awareness (H3). When the questions for this variable were created in the questionnaire, they were designed in line with performance consequences, building a future, and collective activity. Awareness is one of the determining factors for dropping out of school. This can play an important role, but is not always sufficient to prevent dropout, as each decision is influenced by numerous personal and socio-economic contexts, but most people who choose to drop out end up regretting their decision years later. According to a study conducted by Alhassan et al. (2015), which analyzed the experiences of girls who had dropped out of school, 88.3% of all respondents regretted dropping out. According to The New York Times (AP, 1983), in a survey conducted in 1980, 14% of high school students dropped out of high school and later regretted it.

As for the attitude variable (H2), the questions in the questionnaire were created taking into account the awareness of the dropout phenomenon, the feeling of productivity as well as the involvement in reducing dropout. A student's attitude is also closely linked to their personal motivation and interest in school. In the study by Vallerand et al. (1997), they concluded that motivation is a key variable in dropout. They identified four types of motivation to distinguish dropouts from persistent dropouts: self-determined, intrinsic, identified and introjected motivation. In addition, students who are not motivated or do not see a connection between education and their future career are more likely to drop out of school. The degree of confidence and commitment that a student or pupil has towards a course of study reflects their level of career decision-making ability (Gordon, 1998; Robbins et al., 2006).

At the gender-specific level, in addition to H2 and H3, which are statistically significant, there is also H1 for male respondents (variable alcohol or substance abuse). With regard to the intention to stop smoking, it can be seen that men are easier to influence than women. Alcohol and drug use is an important factor influencing the intention to drop out, affecting both academic performance and students' psychological and social well-being. According to one study of university students, drug and alcohol use increased significantly more among males (Wagner et al., 2007). Several analyses have

been conducted over time that found alcohol and drug use to be a significant factor in intent to drop out of school (O'Hara et al., 1998; Fernández-Suárez et al., 2016).

Identifying these important factors is a very important step in preventing dropout. Although it is necessary to prioritize those that have greater importance, it is not advisable to neglect those that have gradually dropped out of the model construct. Knowing these factors, both teachers and parents of students can intervene when they see inappropriate behavior to change the student's decision. As far as dropping out of school is concerned, it is very important to be aware of the consequences that it entails. These include: Restriction of access to well-paid and stable jobs, which affects income and quality of life, loss of important opportunities for personal and professional development, adaptation to changes in society and the demands of the labor market, which are much more difficult to achieve as a result, negative effects on mental and physical health due to vulnerability to financial stress and unhealthy lifestyles, and restriction of the ability to interact and actively participate in the community. Keeping students constantly motivated is key to keeping them in close contact with the academic environment and coping with any external temptations that arise along the way.

Taking into account the results obtained by other researchers, Kumar et al. (2017) carried out a similar work to the one we conducted, also starting from the most frequently identified influencing factors in the literature, to observe the factors influencing school dropout. They found among the variables that have a great influence, similar to our analysis: family structure, parental qualifications, parental occupation, drug addiction or school rules. In a study by Janosz et al. (1997), family, school, behavioral, personal and social variables were identified as predictors of school dropout. In contrast to our significant factors, a stronger influence of school experience variables was found in this work, and the contribution of psychosocial variables, although significant, did not noticeably improve the model. In another study (Pittman, 1991), variables related to the social relationships a student develops within the institution, sense of belonging, peer interest, or teacher-student relationships were found to be strongly associated with school dropout. Many of the variables that were identified as significant in other papers were also variables that were analyzed in our work, even if they only emerged as optimal parameters when the final model was created. These results are influenced by the database used, in particular the respondents and their preferences.

5. Conclusions

Dropping out of school was and still is one of the most important issues to be discussed and, above all, researched in order to find as many appropriate methods as possible to combat it by identifying the factors that lead to its implementation and to discover the factors that make the pupil/student continue his/her studies.

Based on the variables absenteeism, alcohol or substance abuse, attitude, awareness, family, family supervision, school environment and school rules, which were determined by reading various works on the subject of dropping out of school, the influence of these variables on the intention to drop out of school was investigated. For data collection, a questionnaire was created in Google Forms containing a total of 28 questions to analyze each of the above variables. The questionnaire was answered by a total of 669 people.

As part of the analysis, a model was created by processing the data with the SmartPLS program and using the PLS-SEM and Bootstrapping algorithm, with the main objective of observing the variables that influence school dropout. Thus, the greatest influence was found for the variables alcohol or substance abuse, attitude and awareness. In terms of statistical significance, both the initial model and a gender-specific model were tested. It was found that only the alcohol or substance abuse variable was not semi-significant at the level of the overall model, as in the case of the female model.

For the continuity of this analysis, further variables could be identified, also based on the literature, to create new questions to be included in the initial questionnaire. In addition, the number of respondents could be larger and include respondents from different social classes, age categories or backgrounds in order to be able to make any comparison with the initial analysis.

Acknowledgments

This work was funded by the EU's NextGenerationEU instrument through the National Recovery and Resilience Plan of Romania - Pillar III-C9-I8, managed by the Ministry of Research, Innovation and Digitization, within the project entitled „Place-based Economic Policy in EU's Periphery – fundamental research in collaborative development and local resilience. Projections for Romania and Moldova (PEPER)", contract no. 760045/23.05.2023, code CF 275/30.11.2022. This study was co-financed by The Bucharest University of Economic Studies during the PhD program.

References

Afoakwah, C. and Koomson, I. (2021). How does school travel time impact children's learning outcomes in a developing country? *Review of Economics of the Household*, 19(4), pp. 1077-1097. https://doi.org/10.1007/s11150-020-09533-8.

Alhassan, J. (2015). Challenges Faced by Girls who Dropout from Senior High School: the Voices of Sixty Dropout Girls in the Upper West Region of Ghana. *International Journal of Humanities and Social Science*, 5(6), pp. 150-160.

AP. (1983). Survey finds most dropouts have regrets. *The New York Times*. [online] 11 Dec. Available at: <https://www.nytimes.com/1983/12/11/us/survey-finds-most-dropouts-have-regrets.html> [Accessed 8 September 2024].

Berens, J., Schneider, K., Görtz, S., Oster, S. and Burghoff, J. (2018). *Early Detection of Students at Risk – Predicting Student Dropouts Using Administrative Student Data and Machine Learning Methods*. https://doi.org/10.2139/ssrn.3275433.

Bradshaw, C.P., O'Brennan, L.M. and McNeely, C.A. (2008). Core competencies and the prevention of school failure and early school leaving. *New Directions for Child and Adolescent Development*, 2008(122), pp. 19-32. https://doi.org/10.1002/cd.226.

Bridgeland, J.M., DiIulio, J.J. and Morison, K.B. (2006). *The Silent Epidemic: Perspectives of High School Dropouts*. [online] Civic Enterprises. Available at: <https://eric.ed.gov/?id=ED513444> [Accessed 8 September 2024].

Cheah, J.-H., Sarstedt, M., Ringle, C., Ramayah, T. and Ting, H. (2018). Convergent validity assessment of formatively measured constructs in PLS-SEM: On using single-item versus multi-item measures in redundancy analyses. *International Journal of Contemporary Hospitality Management*, 30(11), pp. 3192-3210. https://doi.org/10.1108/IJCHM-10-2017-0649.

Chin, W., Marcolin, B. and Newsted, P. (2003). A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption Study. *Information Systems Research*, 14(2), pp. 189-217. https://doi.org/10.1287/isre.14.2.189.16018.

De Witte, K., Cabus, S., Thyssen, G., Groot, W. and van den Brink, H.M. (2013). A critical review of the literature on school dropout. *Educational Research Review*, 10, pp. 13-28. https://doi.org/10.1016/j.edurev.2013.05.002.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P. and Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. *Journal of the Academy of Marketing Science*, 40(3), pp.434–449. https://doi.org/10.1007/s11747-011-0300-3.

Dijkstra, T.K. and Henseler, J. (2015). Consistent Partial Least Squares Path Modeling. *MIS Quarterly*, 39(2), pp. 297-316. https://doi.org/10.25300/MISQ/2015/39.2.02.

Ding, P. and Feng, S. (2022). How School Travel Affects Children's Psychological Well-Being and Academic Achievement in China. *International Journal of Environmental Research and Public Health*, 19(21), | p. 13881. https://doi.org/10.3390/ijerph192113881.

Esch, P., Bocquet, V., Pull, C., Couffignal, S., Lehnert, T., Graas, M., Fond-Harmant, L. and Ansseau, M. (2014). The downward spiral of mental disorders and educational attainment: a systematic review on early school leaving. *BMC Psychiatry*, 14(1), p. 237. https://doi.org/10.1186/s12888-014-0237-4.

Fernández-Suárez, A., Herrero, J., Pérez, B., Juarros-Basterretxea, J. and Rodríguez-Díaz, F.J. (2016). Risk Factors for School Dropout in a Sample of Juvenile Offenders. *Frontiers in Psychology*, 7, p. 1993. https://doi.org/10.3389/fpsyg.2016.01993.

Fernando, J. (2024). *R-Squared: Definition, Formula, Uses, and Limitations*. [online] Investopedia. Available at: <https://www.investopedia.com/terms/r/r-squared.asp> [Accessed 7 September 2024].

Freudenberg, N. and Ruglis, J. (2007). Reframing School Dropout as a Public Health Issue. *Preventing chronic disease*, 4(4), A107.

Gaik, L.P., Abdullah, M.C., Elias, H. and Uli, J. (2010). Development of Antisocial Behaviour. *Procedia - Social and Behavioral Sciences*, 7, pp. 383-388. https://doi.org/10.1016/j.sbspro.2010.10.052.

Gordon, V.N. (1998). Career Decidedness Types: A Literature Review. *The Career Development Quarterly*, 46(4), pp. 386-403. https://doi.org/10.1002/j.2161-0045.1998.tb00715.x.

Grau-Valldosera, J. and Minguillón, J. (2014). Rethinking Dropout in Online Higher Education: The Case of the Universitat Oberta de Catalunya. *International Review of Research in Open and Distributed Learning*, 15(1), pp. 290-308. https://doi.org/10.19173/irrodl.v15i1.1628.

Gubbels, J., van der Put, C.E. and Assink, M. (2019). Risk Factors for School Absenteeism and Dropout: A Meta-Analytic Review. *Journal of Youth and Adolescence*, 48(9), pp. 1637-1667. https://doi.org/10.1007/s10964-019-01072-5.

Haenlein, M. and Kaplan, A. (2004). A Beginner's Guide to Partial Least Squares Analysis. *Understanding Statistics*, 3(4), pp. 283-297. https://doi.org/10.1207/s15328031us0304_4.

Haimovich, F., Vazquez, E. and Adelman, M. (2021). *Scalable early warning systems for school dropout prevention: Evidence from a 4.000-school randomized controlled trial*. [Working Paper] Documento de Trabajo. Available at: <https://www.econstor.eu/handle/10419/250374> [Accessed 8 September 2024].

Hair, J.F., Ringle, C.M. and Sarstedt, M. (2011). PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, 19(2), pp. 139-152. https://doi.org/10.2753/MTP1069-6679190202.

Hair Jr, J.F, Sarstedt, M., Hopkins, L. and G. Kuppelwieser, V. (2014). Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *European Business Review*, 26(2), pp. 106-121. https://doi.org/10.1108/EBR-10-2013-0128.

Hair, J., Hult, G.T.M., Ringle, C. and Sarstedt, M. (2014). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: SAGE Publications.

Hancco-Monrroy, D.E., Caballero-Apaza, L.M., Abarca-Fernández, D., Castagnetto, J.M., Condori-Cardoza, F.A., De-Lama Moran, R., Carhuancho-Aguilar, J.R., Gutierrez, S., Gonzales, M., Berduzco, N., Delgado Bolton, R.C., San-Martín, M. and Vivanco, L. (2024). Medical Professionalism and Its Association with Dropout Intention in Peruvian Medical Students during the COVID-19 Pandemic. *Behavioral Sciences*, 14(8), p. 641. https://doi.org/10.3390/bs14080641.

Hassan, M.A., Muse, A.H. and Nadarajah, S. (2024). Predicting Student Dropout Rates Using Supervised Machine Learning: Insights from the 2022 National Education Accessibility Survey in Somaliland. *Applied Sciences*, 14(17), p. 7593. https://doi.org/10.3390/app14177593.

Janosz, M., LeBlanc, M., Boulerice, B. and Tremblay, R.E. (1997). Disentangling the Weight of School Dropout Predictors: A Test on Two Longitudinal Samples. *Journal of Youth and Adolescence*, 26(6), pp. 733-762. https://doi.org/10.1023/A:1022300826371.

Jöreskog, K.G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36(4), pp. 409-426. https://doi.org/10.1007/BF02291366.

Kogan, S.M., Luo, Z., Murry, V.M. and Brody, G.H. (2005). Risk and protective factors for substance use among African American high school dropouts. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors*, 19(4), pp. 382-391. https://doi.org/10.1037/0893-164X.19.4.382.

Kumar, M., Singh, A. and Handa, D. (2017). Literature Survey on Educational Dropout Prediction. *International Journal of Education and Management Engineering*, 7(2), pp. 8-19. https://doi.org/10.5815/ijeme.2017.02.02.

Lidbe, A., Li, X., Adanu, E.K., Nambisan, S. and Jones, S. (2020). Exploratory analysis of recent trends in school travel mode choices in the U.S. *Transportation Research Interdisciplinary Perspectives*, 6, p. 100146. https://doi.org/10.1016/j.trip.2020.100146.

Mason, C.H. and Perreault, W.D. (1991). Collinearity, Power, and Interpretation of Multiple Regression Analysis. *Journal of Marketing Research*, 28(3), pp. 268-280. https://doi.org/10.2307/3172863.

Molla, M.T., Madans, J.H. and Wagener, D.K. (2004). Differentials in Adult Mortality and Activity Limitation by Years of Education in the United States at the End of the 1990s. *Population and Development Review*, 30(4), pp. 625-646. https://doi.org/10.1111/j.1728-4457.2004.00035.x.

Mooney, C.Z. and Duval, R.D. (1993). *Bootstrapping: A Nonparametric Approach to Statistical Inference*. 1st edition ed. Newbury Park, Calif: SAGE Publications, Inc.

Netemeyer, R.G., Bearden, W.O. and Sharma, S. (2003). *Scaling Procedures: Issues and Applications*. Thousand Oaks, CA: SAGE Publications.

Nik Jaafar, N.R., Tuti Iryani, M.D., Wan Salwina, W.I., Fairuz Nazri, A.R., Kamal, N.A., Prakash, R.J. and Shah, S.A. (2013). Externalizing and internalizing syndromes in relation to school truancy among adolescents in high-risk urban schools. *Asia-Pacific Psychiatry*, 5(S1), pp. 27-34. https://doi.org/ 10.1111/appy.12072.

O'Hara, P., Parris, D., Fichtner, R.R. and Oster, R. (1998). Influence of Alcohol and Drug Use on AIDS Risk Behavior among Youth in Dropout Prevention. *Journal of Drug Education*, 28(2), pp. 159-168. https://doi.org/10.2190/CXEL-WGAA-DHUK-4F02.

Park, S. and Kim, Y. (2016). Prevalence, correlates, and associated psychological problems of substance use in Korean adolescents. *BMC Public Health*, 16, p 79. https://doi.org/10.1186/s12889-016-2731-8.

Patterson, G.R., Capaldi, D.M. and Bank, L. (1991). An early starter model for predicting delinquency. In D.J. Pepler & K.H. Rubin (Eds.), *The Development and Treatment of Childhood Aggression*. Hillsdale, N.Y: Lawrance Erlbaum.

Pittman, R.B. (1991). Social Factors, Enrollment in Vocational/Technical Courses, and High School Dropout Rates. *The Journal of Educational Research*, 84(5), pp. 288-295. https://doi.org/10.1080/ 00220671.1991.10886029.

Ramírez-Hassan, A., García, G.A., Saravia, E., Duque, J.F. and Londoño, D. (2023). What kind of schools parents choose when they have more options? Effects of school transport subsidies. *Socio-Economic Planning Sciences*, 87, p. 101509. https://doi.org/10.1016/j.seps.2023.101509.

Robbins, S.B., Allen, J., Casillas, A., Peterson, C.H. and Le, H. (2006). Unraveling the differential effects of motivational and skills, social, and self-management measures from traditional predictors of college outcomes. *Journal of Educational Psychology*, 98(3), pp. 598-616. https://doi.org/10.1037/0022-0663.98.3.598.

Theodorou, C., Karekla, M. and Panayiotou, G. (2024). Enhancing Emotion Regulation Skills in High-Risk Adolescents Due to the Existence of Psychopathology in the Family: Feasibility and Uncontrolled Pilot Study of a Group Intervention in a Naturalistic School Setting. *International Journal of Environmental Research and Public Health*, 2 (6), p. 738. https://doi.org/10.3390/ijerph21060738.

Vallerand, R.J., Fortier, M.S. and Guay, F. (1997). Self-determination and persistence in a real-life setting: Toward a motivational model of high school dropout. *Journal of Personality and Social Psychology*, 72(5), pp. 1161-1176. https://doi.org/10.1037/0022-3514.72.5.1161.

Wagner, G.A., Stempliuk, V. de A., Zilberman, M.L., Barroso, L.P. and Andrade, A.G. (2007). Alcohol and drug use among university students: gender differences. *Brazilian Journal of Psychiatry*, 29(2), pp. 123-129. https://doi.org/10.1590/S1516-44462006005000033.

Winkleby, M.A., Jatulis, D.E., Frank, E. and Fortmann, S.P. (1992). Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. *American Journal of Public Health*, 82(6), pp. 816-820. https://doi.org/10.2105/ajph.82.6.816.

Appendix A: Research questionnaire

Issue	Acronym	Questions	Total Disagree	Disagree	Neutral	Agree	Total Agree
Attitude	**AT**	School dropout is a serious problem in school (Q1)	□	□	□	□	□
		Learning and going to school makes me feel productive/productive (Q2)	□	□	□	□	□
		I believe that when I support my classmates I contribute to reducing dropout rates. (Q3)	□	□	□	□	□
Alcohol or substance abuse	**AB**	Drinking alcohol/prohibited substances decreases school performance (Q4)	□	□	□	□	□
		Alcohol/prohibited substance dependence is a disruptive factor leading to dropping out of school (Q5)	□	□	□	□	□
		Consumption of alcohol/prohibited substances reduces the ability to think and concentrate, resulting in poorer and poorer school performance, leading to possible dropout (Q6)	□	□	□	□	□
Awareness	**AW**	I am aware that education is important for my future (Q7)	□	□	□	□	□
		I know that if I drop out of school, it will be difficult for me to shape a career (Q8)	□	□	□	□	□
		All members of society should cooperate to solve the problem of school drop-out (Q9)	□	□	□	□	□
Intention to dropout of school	**INT**	I intend to drop out of school if it is sometimes very difficult (Q10)	□	□	□	□	□
		I do not want to continue my studies and actively participate in classes (Q11)	□	□	□	□	□
		If exceptional circumstances arise that will change the current context, I do not intend to complete my studies (Q12)	□	□	□	□	□
Family monitoring	**FM**	I need to be supervised/prompted regularly to keep myself motivated (Q13)	□	□	□	□	□
		When I am not monitored I tend to neglect school (Q14)	□	□	□	□	□
		Parental neglect of my school situation can lead me to drop out (Q15)	□	□	□	□	□
Family	**FAM**	Unpleasant family events (quarrels, divorce) influence my school performance (Q16)	□	□	□	□	□
		Family example contributes to my mindset about learning (Q17)	□	□	□	□	□
		My family's social status has an impact on my school performance (Q18)	□	□	□	□	□

Issue	Acronym	Questions	Total Disagree	Disagree	Neutral	Agree	Total Agree
School environment	**SE**	I often refuse to go to school after a conflict with my classmates (Q19)	□	□	□	□	□
		An argument with one of my teachers makes me not want to attend classes (Q20)	□	□	□	□	□
		Not fitting in the school group makes me want to give up my studies (Q21)	□	□	□	□	□
		Involvement in school and extracurricular activities reduces the risk of dropping out (Q22)	□	□	□	□	□
Absenteeism	**ABS**	A high number of absences is a first step towards dropping out (Q23)	□	□	□	□	□
		I miss more and more unexcused absences because I am thinking of dropping out (Q24)	□	□	□	□	□
		By being absent I reduce my contact with the subject and my classmates, so I feel that I am gradually detaching myself from the school environment (Q25)	□	□	□	□	□
School rules	**SR**	By not having strict school rules, the process of dropping out becomes easy for me (Q26)	□	□	□	□	□
		The very busy school program makes me drop out (Q27)	□	□	□	□	□
		The mismatch between textbooks and syllabus reduces my possibility to study individually and increase my performance, which demotivates me (Q28)	□	□	□	□	□