MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2024, 20(9), em2502
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/15024

Predictive models, as an idea, to advance the secondary to tertiary transition in science courses

Vaughan van Appel [1*], Estherna Pretorius [2], Rina Durandt [3]

[1] Department of Statistics, University of Johannesburg, Johannesburg, SOUTH AFRICA
[2] Department of Botany and Plant Biotechnology, University of Johannesburg, Johannesburg, SOUTH AFRICA
[3] Mathematics Education Division, School of Education, University of the Witwatersrand, Johannesburg, SOUTH AFRICA

Received 11 December 2023 ▪ Accepted 25 July 2024

Abstract

Investigating the transition between the secondary and the tertiary levels is a main theme in mathematics and science education. More so, this paper considers the transition that intersects with the after-effects of COVID-19, or the transition together with an educational context dominated by sociocultural differences and educational disadvantages. With this knowledge in mind, we investigated the effects of predictive mathematical models (multiple regression, logistic regression, and decision trees) to predict *at-risk* students at three time intervals (weeks one, three, and seven) in the semester. The idea was implemented with a first-year life science class of 130 students. Variables from an academic readiness questionnaire along with early assessment grades were used to build these models. Through a Monte Carlo cross validation method, the performance of the executed predictive models was assessed, and limitations were reported. We argue that the results obtained from predictive models can support both lecturers and students in the transition phase. The idea can be expanded to other courses in STEM fields and other educational contexts.

Keywords: academic readiness, at-risk students, COVID-19, life science students, predictive models, secondary-tertiary transition

INTRODUCTION

The education landscape has changed, perhaps forever, after the arrival of COVID-19, yet it is unclear what impact emergency remote teaching had on students (Engelbrecht et al., 2023). More so, in countries with sociocultural difficulties and pre-existing educational disadvantages, such as South Africa, emergency remote teaching could have had a greater impact on education (for an example from a mathematics education context, see Chirinda et al., 2021; and for an example from a first-year mathematics cohort, see Morton & Durandt, 2023). South Africa is a country with educational inequalities and numerous students enter the tertiary environment underprepared (see results from international studies, such as TIMSS by Mullis et al., 2020). In the post-COVID-19 era, students may face several challenges related to prior learning. The transition between education levels and the difficulties associated with the transition is a main theme in education, which was discussed long before the pandemic (for an example from a German context, see Greefrath et al., 2017; for an example from the South African context, see Jacobs & Pretorius, 2016).

The current cohort of tertiary students experienced emergency remote teaching at the secondary level, for about two years. It is unclear how these students are experiencing the transition to university. In most tertiary classrooms, practices have returned to 'normal', whether to a pre-COVID-19 model or a new post-COVID-19 model. However, some questions arise: Should teachers (lecturers) in the post-COVID-19 era be made aware of risks associated with students? Are the problems from the pre-COVID-19 era also present in the post-COVID-19 era, and if so, have the characteristics of former problems changed? Also, what new and creative ideas can the research offer us to identify and address transition problems?

✉ vvanappel@uj.ac.za **(*Correspondence)** ✉ esthernap@uj.ac.za ✉ rina.durandt@wits.ac.za

Contribution to the literature

- Accurately predicting *at-risk* students is notoriously difficult, where there is a trade-off between *early prediction* and *accuracy*.
- Predicting *at-risk* students during the first week of the academic semester, using only the academic readiness questionnaire (ARQ) as a variable, provides valuable time to assist students, even with its low prediction accuracy.
- Predicting *at-risk* students in week seven of the academic semester, using the marks of a large-scale continuous assessment task as a variable, is sufficiently good, but provides less time to assist students.

The transition from the secondary to the tertiary environment, referred to as the secondary-tertiary transition (STT) in the literature, has been studied for more than two decades and results from several studies have been presented at international conferences (e.g., PME and ICME). The term *transition* can be considered as a period of adjustment before settling into a more stable phase and is often associated with a period of change along with the capability to navigate change (Gale & Parker, 2014; Jansen & Van der Meer, 2012). Students who cannot transit may become disengaged and underperform in the first year of tertiary studies (Sun et al., 2017). For an overview of the literature since 2008, see Di Martino et al. (2023). The STT of students in STEM degrees, particularly with a significant component of mathematics, is a concern, which is supported by the high dropout rates (Geisler & Rolka, 2021). However, difficulties in transition are not only related to cognitive factors and specific subjects (such as mathematics) and can also be influenced by sociocultural and contextual factors, as well as experiences in other subjects. Recently, the affective (e.g., Di Martino & Gregorio, 2019; Geisler & Rolka, 2021) and the sociocultural (e.g., Hernandez-Martinez & Williams, 2013; Lim et al., 2022) perspectives, together with the epistemological and cognitive difficulties students face, give a more holistic view on transition. One example from another subject area is Brazilian first-year nursing students (e.g., Rossato et al., 2024).

The term at risk, in the academic context, refers to students who can be identified as having a higher chance of failing a course (module) and dropping out of university before graduation (Horton, 2015), and is merely the endpoint of a gradual process of disengagement related to several factors (Nouwen & Clycq, 2020). This is largely because at-risk students are faced with challenges and circumstances that negatively impact their ability to navigate through the transition phase (for an example from the South African context, see Jacobs & Pretorius, 2016), and they are more likely to make poor choices regarding their choice of degree, or other decisions that can harm their career. The at-risk students' knowledge, skills, academic ability, and motivation are considerably lower compared to non-at-risk students (e.g., Jacobs & Pretorius, 2016; Ribeiro et al., 2019; Sun et al., 2017). A recent study (compare Glaesser et al., 2024) shows that disengagement (related to behavioral, affective, and cognitive aspects) is an early predictor of students being at risk of dropping out and highlights the need for educators to control classroom dynamics and student composition, particularly in classrooms with diverse student backgrounds. Moreover, students who experience grade retention are steadily more at risk (Nouwen & Clycq, 2020). It is widely understood that at-risk status is associated with first-year students, but this phenomenon was also reported in a third-year mathematics class where cognitive barriers existed (see Braatvedt & Durandt, 2021).

The early identification of at-risk students, particularly in the first year, seems important to develop possibilities for educators to remedy students' academic deficiencies and to increase study success. Several methods can be used for early identification. One idea is a diagnostic or entrance test at the start of the course, which usually focuses on cognitive aspects (for an example from an engineering mathematics class, see Durandt et al., 2021). Another idea is to use mathematical models to identify at-risk students using prior students' grades to forecast current student grades (for an example from a statistics class, see Van Appel & Durandt, 2019).

When building predictive models, one needs to train a model to identify which students are likely to fail (i.e., at-risk students) and which are likely to pass. Ideally, we would want to predict students who may be at risk as early as possible in the first academic semester to provide adequate and targeted support. However, it is known (e.g., Granger & Jeon, 2007; Van Appel & Durandt, 2019) that there is a trade-off between early prediction and the accuracy of forecast models. This trade-off scenario is created when the early prediction (forecast) is based on limited information regarding students' success rate (for example, marks or scores from the first week in the academic semester). This is compared to predicting students' success rate later in the semester (for example, marks or scores at the end of week seven in the academic semester) when more accurate information is available, but this could result in too little time for targeted interventions to have an effect.

The study reported here aimed to identify the effects of different mathematical models to predict at-risk students early in the academic semester. Through this study, we intend to support the STT in a science course while taking into consideration that first-year students from a developing country might even be more at risk in this post-COVID-19 era. We used the example from a first-year life science class to demonstrate our research work (see later section for more information on the sample and why the sample was chosen). A broader aim, which is not particularly investigated in this study but can flow from the study, is to use the data generated by the models to design adequate support for at-risk students.

The research questions are, as follows:

1. What are suitable predictive models to improve the STT for life science student teachers from a developing country; what are the advantages and disadvantages?

2. If the predictive accuracy of these models (identified in 1) is compared at different stages during the academic semester, what are the consequences?

In answering these research questions, we attempt to address a gap in the research, particularly research regarding STT in a developing country in Africa, which is underrepresented compared to global standards (see Di Martino et al., 2023). The paper will report on theoretical aspects related to students in transition and the use of predictive models for educational purposes; the case of the first-year life science class; the selection of suitable predictor variables; technical aspects of the construction of the models; results and discussions at different time intervals of the academic semester; followed by concluding remarks and perspectives.

THEORETICAL PERSPECTIVES

The study has a pragmatic nature, and the following two theoretical perspectives informed our work:

(1) the transition of students from the secondary to the tertiary level and characteristics associated with this process, and

(2) the usefulness of mathematical models in the educational context.

The Transition of Students from the Secondary to the Tertiary Level

Research on STT includes a variety of theoretical frameworks and a clear evolution over time is visible from a purely cognitive approach (compare Tall, 2008) to a more holistic approach, which also includes affective constructs (compare Geisler, 2021).

In this study, we investigated the effect of predictive models at different time intervals; thus, Nicholson's (1990) transition cycle seemed a suitable theoretical foundation to inform the STT. This cycle was originally designed for transitions in the workplace, from an organizational behavior perspective; it is dynamic and can be applied to other real-life situations when a transition is required. Leong et al. (2021) used Nicholson's cycle as a theoretical foundation in their study regarding educators' perspectives on first-year chemistry students' transition from the secondary to the tertiary level. Nicholson's (1990) transition cycle is described by four stages that emphasize both the cognitive and the non-cognitive factors that are involved in student readiness: preparation, encounter, adjustment, and stabilization.

A short description of the phases follows:

1. First, the *preparation phase* occurs at the secondary level before students enter the tertiary environment and provides a foundation for the new tertiary environment. Here students should develop readiness, expectations, and motivation. Globally, the current first-year groups experienced several turbulences in their formal education during this phase, such as emergency remote teaching, and they might be less prepared for tertiary studies than their former counterparts.

2. Second, the *encounter phase* is typically associated with the first few weeks at university when students make sense of the new environment, evaluate their ability to cope, and create new social networks. Typically, during this phase, students are well supported by university services. Not all students enter university with similar readiness characteristics, and this unequal scenario creates challenges for teaching.

3. Next, the *adjustment phase* is associated with the remainder of the first year. This is the period where students adapt to their new environment and adopt the necessary behaviors to fit into the environment. The expectation is that students will become less dependent on the support services offered by the university and the lecturer and demonstrate more self-regulated habits.

4. Last, the *stabilization phase* is where students are fully adapted to the academic context. Some authors refer to the 'equilibrium' phase when students can balance all aspects of the academic environment. The expectation is that second-year students should be in this phase, but that might not be the case for all students. According to the literature (Leong et al., 2021), the time for students to reach stabilization varies greatly in length, and academic success can be used to determine whether stabilization has been reached.

Mathematical Models and Their Relevance and Usage for the Educational Context

A mathematical model is a representation of aspects of an extra-mathematical domain using some mathematical entities and relations between them (Niss & Blum, 2020). For an example of a modelling cycle often used in teaching and learning, see Blum and Leiβ (2007).

In the study presented here, our focus was not to use a mathematical model for teaching or learning, but rather to apply the modelling process to solve a real-life problem. Mathematical models are usable at all levels of society. Some of these are to control processes, to design products, and to monitor and influence systems.

The modelling process is cyclic. The first step is usually to identify the problem in the real world; for example, low throughput rates in a science course. The second step is to make assumptions and identify suitable variables by selecting relevant information and searching for relations; for example, using small-scale and large-scale continuous assessment marks as variables for the model. The third step is to formulate a mathematical model and perform procedures to find results; for example, to use statistical methods such as a linear regression model to find results. The fourth step is to analyze and assess the solution by questioning the results and consequences; for example, to check the model's accuracy and consider if the model is suitable in the context and can be used for prediction. The last step is to iterate the modelling process to refine and extend the model; for example, by using different data sets or adding variables to the model.

It seems reasonable to consider using a mathematical model to predict at-risk students in the STT, but if such a model is used to improve decision-making (e.g., regarding study success or adequate support) then the quality of the model is also important. Thus, in reference to the aim of this study, the emphasis should be on the product and its efficiency, as it would play a larger role in decision-making, rather than on the modelling process itself.

One example to illustrate how the proposed mathematical model could potentially impact decision-making in the context of STT is to monitor student grades at three time intervals in the semester. As mentioned before (see previous section), several characteristics describe at-risk students; one of the characteristics is when students experience a grade retention, which means they are steadily more at risk. Therefore, the continuous monitoring of student progress is imperative to identify students who are at risk of failing a course. The time factor is often found to be an influential factor (see van der Put, 2020), and early prediction of at-risk students can lead to early prevention and early support (for example, assisting students via targeted consultation sessions with the course tutors and lecturer). The time factor is crucial as students can be at risk at any time during the semester (even if they were not at risk at the start of a course), or they might be at risk at the start of a course and become more stable a few weeks later. Thus, in larger first-year classes, results from a suitable predictive model (or models) can help lecturers and course administrators to identify and keep track of at-risk students and plan targeted support at specific time intervals.

During the evaluation of a model, the following criteria should be considered (see James et al., 2021; Meyer, 2012):

(1) accuracy (referring to the correctness of the output values),
(2) realism (referring to the correctness of assumptions),
(3) precision (questioning if its predictions are in exact numbers),
(4) robustness (questioning if the model is to some extent protected against errors in the input data),
(5) generalizability (considering if the model can be applied in another context), and
(6) fruitfulness (questioning if the results bring forward useful conclusions).

Unfortunately, we do not have a single model that can be applied to all situations, and, therefore, we need to evaluate various models to find a suitable model for the problem at hand (see James et al., 2021, chapter 2).

Through predictive modelling, we intend to use mathematical and computational methods to predict students' probability of academic success based on changes in the model input values. A limitation in building accurate predictive models is that these models are largely dependent on the quality of the data available. In this study, the quality and reliability of the data (which were only collected in one academic semester) influence the accuracy (or predictive power) of the models. A true evaluation of the accuracy of the models used will require observations over several years.

DESIGN AND METHODS

The idea to conduct the study developed at the end of the COVID-19 era when business as usual was again widely promoted and several lecturers discussed the unknown effect of the pandemic on education. The study was implemented from February to April 2022. Knowing the challenges associated with predictive mathematical models (such as the quality of data) the research team decided to collect data from participants at three time intervals in the academic semester; in weeks one, three, and seven. These specific intervals were chosen for three reasons–the transition phase of a first-year student, the time factor (aligned with the idea that early identification leads to early prevention), and the alignment with course assessments (both small and large-scale assessments) and the assessment structure.[1]

[1] The assessment structure, to calculate the final course mark (FM), consists of small-scale assessments (class tests, tutorials, and practical) together 10%, two large-scale assessments (semester tests) each 20%, and a final examination of 50%. Six content areas

Data were collected in week 1 via the ARQ, then again in week 3 from results obtained in small-scale continuous assessment tasks, and in week 7 from the results obtained from a large-scale continuous assessment task (see later sub-sections for details). The data were used to calibrate predictive models and to identify suitable models for the context (see later sub-section for details). The study was quantitative-oriented, and data were collected and processed according to the university's standard ethical procedures.

Participants

The participants in this study included 130 student teachers in the first-year life science class from a large public university in Johannesburg. Although the participants are from the Faculty of Education, the course is offered in the Faculty of Science. The sample was conveniently chosen as the second author was the course lecturer who is an experienced lecturer in undergraduate life science courses. Further, the content in the first-year life science course contains several challenging topics that are not normally treated at the school level, which created an ideal context for the study.

Even though English is the language of instruction at the university, most students speak one of the other official languages (e.g., Zulu) at home. South Africa is a country with numerous educational disadvantages, and participants in this study likely had diverse school experiences. Data on participants' school experiences during the pandemic and assessment results at the secondary level are not available. Data collected from the entire cohort was used to build and 'train' the predictive models.

Predictor Variables

Several variables were used to predict at-risk students in the life science course in the first academic semester, at three time intervals. More details on the predictor variables and when they were available throughout the course are outlined below. The reasons for considering variables at the three time intervals (as mentioned earlier) were mainly to continuously monitor students' progress from the start of the course (aligned with the assessment structure) and to more accurately identify/monitor at-risk students using informative variables as students' progress through the course. We selected predictor variables that were fairly easy to obtain and that provided sufficient data, such as an online survey (ARQ) and assessment marks. The natural expectation was that as students progressed through the semester, we would gradually gain more insight into the students' possible success (or failure) in the course. A broader aim following from this research initiative would be to provide adequate and effective support to at-risk students. However, by waiting too long in the academic semester to gather data on students' progress, educators would have less time for meaningful interventions to support struggling students with less promising results. Thus, the time factor was considered in selecting suitable predictive models. It is more difficult for students to make up marks in a cumulative weighted assessment grading system later in an academic semester. Next, we explain the predictor variables considered in the study and the various time intervals:

Week 1–Academic readiness questionnaire

During the first week of the academic semester, we had no information available to build a predictive model for first-year students. Therefore, we decided to collect data from students in week one via an ARQ. The ARQ was available to all students on the University's learning management system, Blackboard. Various surveys that measure the academic readiness of students exist in the literature (e.g., Bitzer, 2003; Lemmens, 2010; Pintrich & De Groot, 1990), but the ARQ developed by Lemmens (2010) seemed more appropriate for this study as it was designed to measure South African students' readiness for tertiary education. The items were sourced from other well-studied surveys, then Lemmens (2010) tested and determined their relevance and comprehensibility in the South African context.

The ARQ consists of 70 questions measured on a Likert scale (1–definitely disagree, 2–disagree, 3–neutral, 4–agree, 5–definitely agree). Through a sample of 1,222 first-year students in the Faculty of Economic and Management Sciences, Lemmens (2010) identified five potential factors in an orthogonal factor analysis with eigenvalues larger than two, which explain a combined total of 31.7% of the cumulative variance. Further comparative analysis was carried out on a reduced three-factor analysis. However, this did not produce better results and ultimately the five factors were retained. The five factors, developed through this analysis process (Lemmens, 2010), to identify academic readiness are labelled:

(1) *achievement motivation orientation*,

(2) *learning efficiency*,

(3) *goal orientation*,

(4) *integration and support*, and

(5) *reading behavior*.

were treated, and each content area was assessed by small-scale assessments. Each semester test focused on three content areas and the examination focused on all content areas.

Table 1. The five factors of the academic readiness questionnaire (Lemmens, 2010, p. 161)

Factor name	Definition
Achievement motivation orientation (17 questions)	The degree to which one has an intrinsic interest in higher education and an expectation to achieve academically. *Example statement: I can be successful in my studies this year.*
Learning-efficacy (12 questions)	The degree of confidence in one's own ability to achieve one's academic goals. *Example statement: I have confidence in sharing my own opinions, even if they might be different from the way most other people think.*
Goal orientation (11 questions)	The degree to which one can plan for learning by setting task-specific goals. *Example statement: I set specific goals before I begin learning for tests/exams.*
Integration and support (14 questions)	The degree to which student experiences institutional, social, family, and financial support. *Example statement: If I run into problems at university, I have someone who would help me.*
Reading behavior (6 questions)	The degree to which one enjoys reading for pleasure. *Example statement: I will try to do optional reading even though I know it will not influence my performance.*

Table 2. Descriptive statistics for the five factors from the ARQ

	Achievement	Efficacy	Goal orientation	Integration	Reading
Mean	4.2	3.6	3.7	3.5	3.8
Median	4.3	3.7	3.7	3.5	3.8
Mode	4.6	3.8	3.8	3.4	4.5
Standard deviation	0.6	0.6	0.5	0.5	0.7
Kurtosis	5.0	-0.1	0.0	0.0	-0.5
Skewness	-1.9	-0.5	-0.1	-0.3	-0.3
Minimum	1.7	2.0	2.5	2.2	2.2
Maximum	4.9	4.7	5.0	4.5	5.0
Count	88.0	88.0	88.0	88.0	88.0

Low scores for the first four categories are associated with risk, whereas high scores for *reading behavior* are associated with risk. Further explanations of the five factors, along with an example question from each factor, are given in **Table 1**.

Determining whether high reading behavior scores are associated with at risk is peculiar to the study and would need further investigation. However, it is likely that reading ability and comprehension would typically be a stronger measure related to academic readiness than reading behavior (Du Plessis et al., 2005; Lemmens, 2010, p. 249), but it is also more difficult to measure. Further, the questions related to the reading behavior in the ARQ do not inspire student confidence in the course content, which would support the association between high scores and being at risk. This is evident in the example question given in **Table 1**.

The ARQ was voluntary for students, and 88 responses were recorded (out of a class of 130 students). For each student, we calculated their mean Likert score for the five factors shown in **Table 1**, representing the student's mean score for each factor. Descriptive statistics were then calculated on these scores and shown in **Table 2**. The descriptive statistics show that most students are ready for tertiary education (sample mean, median, and mode scores are between 3.4 and 4.5 for each category out of a total score of 5 with a low standard deviation in the sample data), except for the category *reading behavior*, where high scores are associated with being at risk.

Any model built using the ARQ data can only forecast a future student's expected outcome if they completed the questionnaire during week 1. Although the ARQ showed promising results for students, we expected lower mean scores.

Week 3–Small-scale assessment tasks (class test 1, tutorial 1, and practical 1)

By week three of the academic semester, we have accumulated further data in the form of students' marks from three small-scale assessment tasks (*class test 1*, *tutorial 1*, and *practical 1*). All three assessment tasks focused on the content in the biochemistry section of the life science course (i.e., the macromolecules of life: carbohydrates, proteins, lipids, and nucleic acids) that was presented during the first three weeks of the semester. A class test is an individual assessment of an entire section (e.g., *class test 1* on the macromolecules of life) administered in official class time. A tutorial task is an in-class individual or group activity on the content covered in one week during a tutorial session (e.g., *tutorial 1* on carbohydrates), and students can be supported by a course tutor if needed. A practical task is an individual hands-on activity in a laboratory where students apply the theory covered in each section (e.g., *practical 1* on carbohydrates, proteins, lipids, and nucleic acids) and laboratory assistants provide support, related to practical matters, to students, if needed. From week one to three, the lecturer (and the second author of this paper) monitored class attendance carefully and

Table 3. Descriptive statistics of small-scale assessment tasks (*class test 1*, *tutorial 1*, and *practical 1*) and a large-scale assessment task (*semester test 1*)

	Week 3 (contributing 5% of FM)			Week 7 (contributing 20% of FM)
	Class test 1	*Tutorial 1*	*Practical 1*	*Semester test 1*
Mean	56.7	66.3	54.1	64.1
Median	61.0	69.0	56.0	65.0
Mode	0.0	68.0	58.0	68.0
Standard deviation	23.5	19.9	15.7	12.5
Kurtosis	0.7	4.3	1.3	1.0
Skewness	-1.1	-1.9	-0.7	-0.6
Minimum	0.0	0.0	0.0	22.0
Maximum	97.0	99.0	90.0	92.0
Count	130.0	130.0	130.0	130.0

Table 4. Correlation between test scores, both small-scale, and large-scale assessments

	Class test 1	*Tutorial 1*	*Practical 1*	*Semester test 1*	FM
Class Test 1	1.0000				
Tutorial 1	0.6184	1.0000			
Practical 1	0.3450	0.3813	1.0000		
Semester test 1	0.4574	0.4853	0.3493	1.0000	
FM	0.4718	0.4652	0.4144	0.7396	1.0000

controlled the implementation of all learning situations and assessments. The idea was to keep a close eye on student progress and performance and to ensure all learning activities ran smoothly. According to the assessment structure, these three small-scale assessments contributed 5% towards the FM. The assessments were compulsory and provided a complete dataset of 130 responses. We expected the data could give further insight into students' future success rate in the course and included this in the construction of the models (see later section for details). Descriptive statistics are shown in **Table 3**.

It is evident from the small-scale assessment marks, accumulated by week three, that most students are proceeding well (mean marks greater than 50% indicate a mean pass mark for each assessment). It might be that students at this stage, while they experienced substantial lecturer support and could still master the volume of content, did not show cracks. We argue that by predicting at-risk students early in the semester, even if they do not yet show substantial evidence of an at-risk status (e.g., failing an assessment), educators can provide additional support to lower-performing students who might fall behind later in the semester.

Week 7–Large-scale assessment task (semester test 1)

During weeks four to six of the academic semester, students were exposed to two new sections (on the topics of cells and mitosis) during lectures, as well as further small-scale assessments. In week seven, a large-scale summative assessment was scheduled on content sections one to three, which contributes 20% towards the FM. Aside from the final examination, this is one of two large-scale assessments in the course. A summary of descriptive statistics on the data collected from *semester test 1* in week seven is shown in **Table 3**. Similar statistical measures of center and spread are found in the large-scale assessment and the accumulated small-scale assessments. Further, the linear correlation between the test scores and the FM was calculated and is shown in **Table 4**. Correlation measures the strength of the relationship between two variables, which is important when using information from one variable to predict another variable's outcome.

Results show a strong positive linear correlation between the large-scale assessment (*semester test 1*) and the FM (r = 0.7396), which suggests that predictive models based on linear relationships between variables are likely to be good models for identifying at-risk students. It is further evident from **Table 3** that as the semester progresses, we obtain more insight into the expected success of students. Since all assessment variables are positively correlated with the FM, it follows naturally to start predicting at-risk students early, even with lower accuracy, but with the broader idea of targeted support for struggling students.

Predictive Models

In this study, we used three predictive models, namely a *multiple linear regression model*, a *logistic regression model*, and a *decision tree model*. These models are commonly used in statistical and data mining practices as they are easy to implement in most statistical software packages and have proven to be robust and relatively easy to interpret (see, e.g., James et al., 2021). The practice of data mining is to use statistical techniques to extract patterns and trends in data sets, which can be used to make informative decisions. In this study, we aimed to identify patterns and trends (by statistical models) from the data obtained on students'

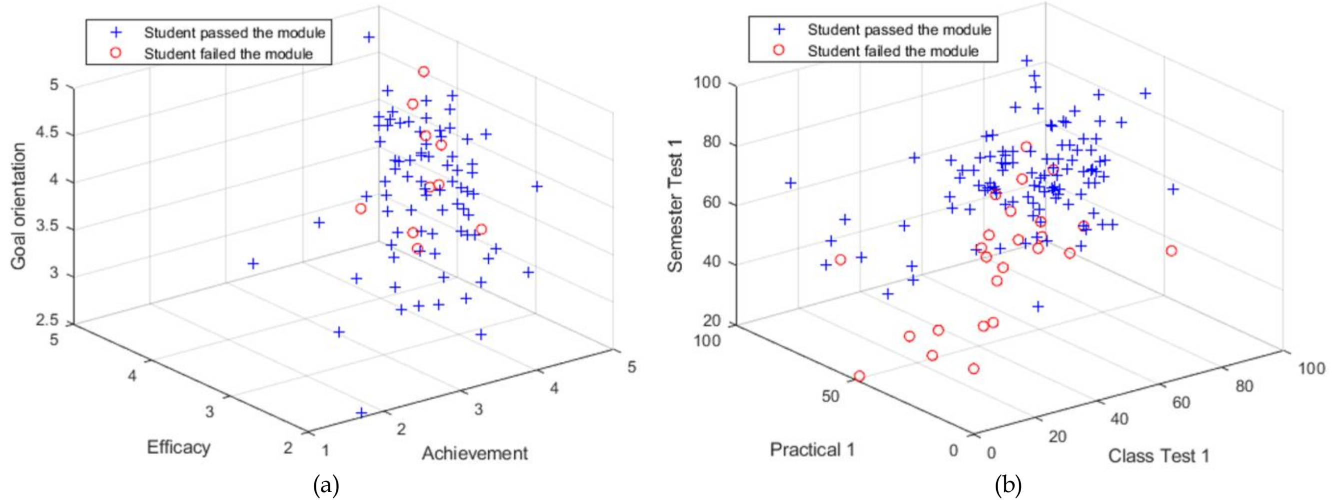

Figure 1. A visual relationship between the covariates and the course outcome in (a) week 1 & (b) week 7 (Source: Authors' own elaboration)

results for the assessments and the ARQ, and we used the models to identify the current students' trajectory. All predictive models were implemented in MATLAB using the built-in statistical toolbox. Technical and mathematical aspects of the predictive models are shared in this section, which might not be interesting to all readers, but we argue it is necessary to explain these aspects to advance the understanding of these models to readers and for possible duplication in other contexts.

In part a in **Figure 1**, the concept of data mining and predictive modelling is illustrated. In particular, **Figure 1** shows the relationship between three covariates obtained in week one from the ARQ questionnaire, namely *achievement*, *efficacy*, and *goal orientation* (presented on the x, y, and z-*axis*), along with the students' outcome in the course (pass $+$/fail $\circ$). Similarly, part b in **Figure 1** shows the relationship between three covariates accumulated by week seven, namely *class test 1*, *practical 1*, and *semester test 1*. We, therefore, aim to use the sample data (as shown in **Figure 1**) to train a model that uses the covariates as an input to identify patterns and trends, which can then be used to predict future students' outcomes (pass or fail) in the course.

In building predictive models, we are interested in understanding the association between the response variable (course outcome), Y, and the available n covariates (predictor variables), $X_1, X_2, \dots, X_n$. To do this, we need to identify the following (see, e.g., James et al., 2021):

1. Which predictor variables are associated with the response variable?
2. What is the relationship between the response and the predictor variables?
3. If there is a relationship, how can the relationship between the response and predictor variable be defined (e.g., as linear)?

Thus, in this study, the goal was to identify which predictor variables–of the five ARQ factors (*achievement motivation orientation*, *learning efficiency*, *goal orientation*, *integration and support*, and *reading behavior*), the three small-scale assessments (*class test 1*, *tutorial 1*, and *practical 1*), and the large-scale assessment (*semester test 1*)–are associated with the response variable (course outcome), and which model is best to capture this relationship. This was done by carrying out model diagnostics after fitting each model to the sample data (i.e., analyzing the model residuals and carrying out goodness-of-fit tests).

Multiple linear regression model

The first model considered in this study assumes a linear relationship between the response variable and the predictor variable. The standard multiple linear regression model is given as a linear relationship between the outcome and the p predictor variables (see, James et al., 2021):

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p + \varepsilon, \tag{1}$$

where X_j represents the j^{th} predictor variable, β_j represents the coefficient that quantifies the association between the predictor variable and the response variable Y, and ε represents a random error term with mean zero and constant variance. Once the model has been fit to the observed data, the relationship between the response variable and predictor variables must be tested (model diagnostics). This is typically done by testing the null hypothesis: $H_0 : \beta_1 = \beta_2 = \cdots = \beta_p = 0$ versus the alternative hypothesis: $H_1\!: at\ least\ one\ \beta_j$ is *non-zero*, and assessing the model residuals. If the null hypothesis is not rejected, it implies that the j^{th} covariate provides no significant information to predict the expected outcome of the student and the variable should be removed from the model.

In this study, standard theoretical practices were followed, and we refer the interested reader to James et al. (2021) for more details regarding the estimation of the parameters, β_j, and the statistical fit of the model. However, most statistical packages can automatically carry out the estimation and produce the model fit analysis.

Logistic regression

The second model is called logistic regression, which is known to be a robust classification model for predicting at-risk students (see e.g., Marbouti et al., 2016; Van Appel & Durandt, 2019).

Logistic regression is a powerful statistical forecast model that is widely used in Binomial data and can be easily implemented in many statistical programs. Rather than forecasting the students' grade (Y) directly, logistic regression models predict the probability that the final grade (Y) belongs to a certain category (i.e., 0 or 1) (see, James et al., 2021). For example, the probability of passing the course given a set (vector) of predictor variables, X, is denoted as $p(X) = P(Y = 1|X)$. We, therefore, model the function $p(X)$ using the following logistic function (James et al., 2021):

$$p(X) = \frac{e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}}{1+e^{\beta_0+\beta_1 X_1+\cdots+\beta_p X_p}}. \qquad (2)$$

Eq. (2) can then be re-expressed as the log odds of passing the course:

$$log\left(\frac{p(X)}{1-p(X)}\right) = \beta_0 + \beta_1 X_1 + \cdots + \beta_p X_p. \qquad (3)$$

The logistic regression has *log* odds that are linearly related to X, where small values for the log odds represent high chances of failing the course. Standard theoretical practices, which are similar to applying the multiple linear regression model, were followed, and after fitting the logistic regression model to the sample data we carried out standard residual diagnostics and goodness-of-fit tests to assess the appropriateness of the model for the problem.

Decision trees

The third method is tree-based, which can be applied to both regression and classification problems. Tree-based methods have the general advantage that they are simple to interpret and involve stratifying the predictor space into several simple paths (see, e.g., James et al., 2021, chapter 8). **Figure 2** and **Figure 3** present visual representations of this idea for two decision trees. With reference to this study, variable x_1 represents the mark for *class test 1*, variable x_2 represents the mark for *tutorial 1*, and variable x_3 represents the mark for *practical 1*. Furthermore, for the classification model, 0 represents a 'fail', and 1 a 'pass'. At each node (triangle) of the tree, a question is asked, where a decision is made to branch to the left or the right. Each path is assigned to a terminal node (denoted as the dot) based on the answers to the questions. The forecast results are indicated below each terminal node.

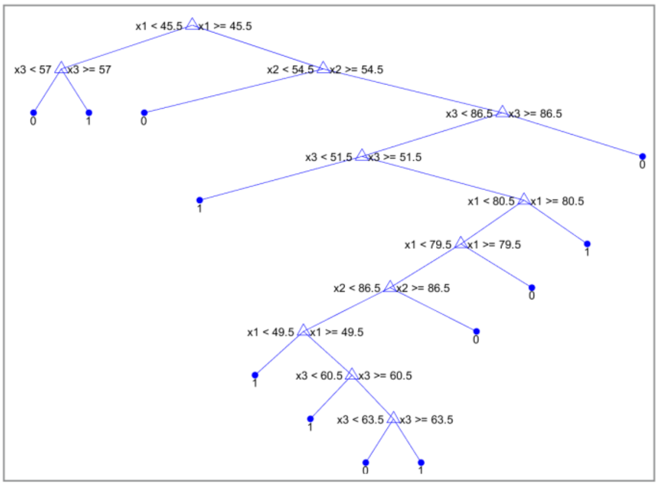

Figure 2. Classification tree with reference to the marks for *class test 1*, *tutorial 1*, and *practical 1* (Source: Authors' own elaboration)

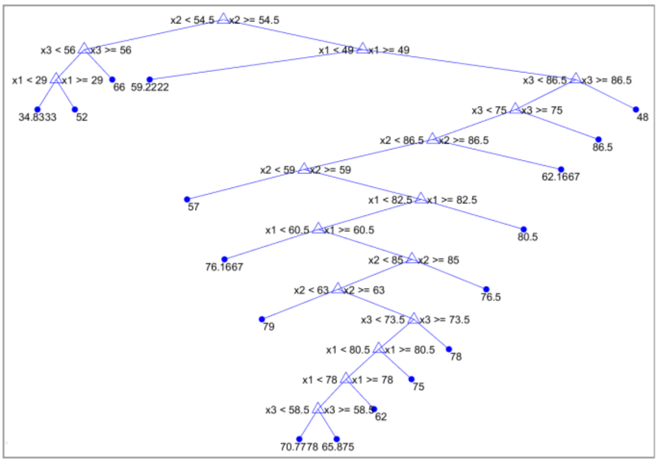

Figure 3. Regression tree with reference to the marks for *class test 1*, *tutorial 1*, and *practical 1* (Source: Authors' own elaboration)

This tree-based process may produce good results on a training set, but it is likely to overfit the data, which will lead to poor test set results. Therefore, there is a need to prune the tree to fewer splits. This will likely lead to lower variance with better interpretations (see James et al., 2021, chapter 8). To find an optimal number of splits, which produce the lowest test error rate, we will make use of a technique called Monte Carlo cross validation (MCCV) (see the next section for clarification).

Monte Carlo cross validation

MCCV is a resampling approach that is often used to evaluate machine learning models (Shan, 2022). This is done by repeatedly selecting training and test splits from the sample data (Kuhn & Johnson, 2016). The researcher will typically decide on the proportion of the sample data that will form the training and test subsets as well as the number of repetitions. In this study, we followed standard theoretical practices and randomly selected, without replacement, 50% of the sample data for training and the remaining 50% for testing, which was repeated 1,000 times. For each of the 1,000 iterations, we calculated the measures of precision and recall (defined in the next section) and reported the average over the 1,000 iterations. The use of cross validation techniques is particularly useful when limited statistical tests are available, such as in the multiple regression and logistic regression models (James et al., 2021, chapter 5). The use of cross validation methods is also useful in deciding on optimally pruning a decision tree (Efron & Tibshirani 1994, chapter 17). The MCCV was used in assessing the accuracy of all the models used in this study.

MODEL RESULTS AND DISCUSSION

In this section, we report on the results, which follow the process of analysis: First, we applied the different predictive models (*multiple linear regression*, *logistic regression*, and *decision trees*) to our sample data and, second, we analyzed the performance and comparison of these models by calculating measures of precision and recall (this is the process to assess and compare models to identify the most suitable predictive model for the problem).

Below is the measure that we used to assess the accuracy of the models. These measures are typically referred to in the literature as measures of *precision and recall* (see, e.g., James et al., 2021, chapter 4; Marbouti et al., 2016):

1. *Harmonic mean:* The harmonic mean ($F_{1.5}$) represents a weighted mean where the accuracy of predicting at-risk students is weighed more than the accuracy of predicting students who are not at risk.
2. *Accuracy:* Proportion of correctly identified students out of the sample.
3. *Accuracy-pass:* Proportion of students correctly identified as not at risk out of the true number of not at-risk students.
4. *Accuracy-fail:* Proportion of students correctly identified as at risk out of the true number of at-risk students.
5. *True negative:* Correctly identified a student as not at risk, and the student passed the course.
6. *True positive:* Correctly identified as at risk, and the student failed the course.
7. *False negative:* The proportion of students who failed the course, where the model incorrectly identified the students as not at risk of failing the course.
8. *False positive:* The proportion of students who passed the course, where the model incorrectly identified the students as at risk of failing the course.

The precision and recall formulas for the above measure are given as (see Marbouti et al., 2016):

$$Accuracy = \frac{True\ Negatives + True\ Positives}{Total\ number\ of\ students}. \quad (4)$$

$$Accuracy_{pass} = \frac{True\ Negatives}{Number\ of\ passed\ students} = \frac{True\ Negatives}{True\ Negatives + False\ Positives}. \quad (5)$$

$$Accuracy_{Fail} = \frac{True\ Positives}{Number\ of\ failed\ students} = \frac{True\ Positives}{True\ Positives + False\ Negatives}. \quad (6)$$

$$F_{1.5} = \frac{(1+1.5^2) \times TP}{(1+1.5^2) \times TP + 1.5^2 \times FN + FP}, \quad (7)$$

where TP is true positives, FN is false negatives, and FP is false positives.

The harmonic mean is particularly useful in this study as there is a larger consequence (and therefore weighed more in the harmonic mean) for not accurately predicting students who are at risk of failing the course, as they would not receive assistance, than accurately predicting students who pass the course.

Week 1–Forecast Results

Both the *multiple linear regression* and *logistic regression* models were inadequate to model at-risk students (based on the residual analysis and goodness-of-fit tests to the sample data) during the first week of the academic semester, by using the ARQ covariates (see earlier section for details). This is due to a non-linear relationship between the covariates and the response variable.

The forecast accuracy results for the *decision tree* models using the ARQ as the predictor variable are shown in **Table 5**. This represents the accuracy of predicting at-risk students during the first week of the semester. Further, we give the standard error (SE) in our accuracy estimates. The SE is a measure of the variation in the estimates (i.e., how accurate the estimates are) over the 1,000 MCCV simulations. We focused on achieving a high *accuracy-fail* and a low *false negative* percentage, with low variation in our estimates (low SE). We regard the model outcome to classify a student as not being at risk, when the student is at risk of failing the course (*false negative*), as problematic with a large consequence for the student.

We can see from **Table 5** that predicting at-risk students during the first week of the semester using the ARQ does not provide a good prediction of at-risk students. Both *decision tree* models achieve a low accuracy in forecasting the students who failed the course with an average harmonic mean of 11-13% with a high (15-16%) forecast SE. Both models reported unsatisfactory *false negative* percentages (10%). Early prediction of at-risk students is not particularly accurate;

Table 5. Cross-verification results using ARQ scores from week 1

Method	Decision tree classification	Decision tree R
Harmonic mean ($F_{1.5}$)	13%	11%
SE	15%	16%
Accuracy	83%	85%
SE	6%	6%
Accuracy-pass	92%	95%
SE	8%	6%
Accuracy-fail	15%	11%
SE	20%	17%
True negative*	81%	84%
SE	7%	6%
True positive*	2%	1%
SE	2%	2%
False negative*	10%	10%
SE	4%	4%
False positive*	7%	5%
SE	7%	6%

Note. *The percentage of students

Table 6. Cross-verification scores using small-scale assessments from week 3 (*class test 1*, *tutorial 1*, and *practical 1*)

Method	Decision tree C	Decision tree R
Harmonic mean ($F_{1.5}$)	33%	34%
SE	15%	13%
Accuracy	77%	77%
SE	5%	5%
Accuracy-pass	88%	88%
SE	7%	7%
Accuracy-fail	33%	33%
SE	17%	16%
True negative*	71%	71%
SE	6%	6%
True positive*	6%	6%
SE	3%	3%
False negative*	13%	13%
SE	5%	4%
False positive*	10%	10%
SE	6%	6%

Note. *The percentage of students

Table 7. Cross-verification scores using a large-scale assessment from week 7 (*semester test 1*)

Method	Logistic regression	Multiple regression	Decision tree C	Decision tree R
Harmonic mean ($F_{1.5}$)	57%	64%	58%	58%
SE	12%	9%	13%	14%
Accuracy	88%	89%	86%	87%
SE	3%	3%	4%	4%
Accuracy-pass	97%	96%	93%	94%
SE	2%	2%	5%	4%
Accuracy-fail	52%	60%	56%	55%
SE	13%	10%	15%	16%
True negative*	78%	78%	75%	76%
SE	4%	3%	5%	4%
True positive*	10%	11%	11%	11%
SE	3%	3%	3%	3%
False negative*	9%	8%	9%	9%
SE	4%	3%	4%	4%
False positive*	3%	3%	6%	5%
SE	2%	2%	4%	3%

however, it identifies students who may require targeted support early in the academic semester. If a student was incorrectly identified as at risk, the student would still benefit from the intervention. Students can be reassessed during week three, with potentially better accuracy. We, therefore, argue that both at-risk and non-at-risk students can benefit from early prediction.

Week 3–Forecast Results

Table 6 shows the accuracy of predicting at-risk students during the third week of the academic semester. Similar to the week one predictions, the *multiple linear regression* and *logistic regression* models were not adequate models for the data. Using only the three small-scale assessments (*class test 1*, *tutorial 1*, and *practical 1*) produced the strongest models during week three. Both *decision tree* models achieved a significant increase in accuracy with an overall harmonic mean accuracy of 33-34% and an *accuracy-fail* score of 33%. Further, we have seen a decrease in the forecasts SE (from lying between 17-20% in week 1 to 16-17% in week 3). Although we can identify an improved accuracy in predicting at-risk students in week three compared to week one, the results are not yet sufficiently good. This is further supported by an increase in both the proportion of *false negative* and *false positive* cases in week three.

Week 7–Forecast Results

In **Table 7**, we show the predictive results for the models using only the marks obtained from the large-scale assessment (*semester test 1*), which were implemented during week seven of the academic semester. Using only the large-scale assessment mark yielded the strongest model by our model diagnostics and cross-validation analysis. However, all models (*multiple* and *logistic regression* and *decision trees*) are viable models. The linear regression models performed well this time due to the strong linear relationship between the mark for *semester test 1* and the FM (see **Table 4**). In **Table 7**, we report an average *harmonic mean* of between 57-64% (with lower forecast SE, between 9-14%) for all models, an overall *accuracy* between 86-89%, and an *accuracy-fail* (correctly predicting a student to fail) of between 52-60%. The forecasts during week seven also yielded the lowest proportion of *false negative* and *false positive* outcomes (between 8-9% and 3-6%, respectively).

Accurately predicting students who are at risk of failing a course is notoriously difficult. Several aspects should be considered, such as suitable variables, models, and time intervals. Concerning the research questions (see earlier section), we can confirm that multiple regression, logistic regression, and decision tree models can be used to predict at-risk students. Our assessment shows that decision tree models are most suitable for predictions in weeks one and three, albeit with limited accuracy. In week 7, all models can be used. However, the most accurate and sufficiently good forecasts can be generated by the multiple linear regression model (highest *accuracy-fail*, lowest *false negative,* and lowest SE percentages). Continuously modelling students' expected outcomes is an idea to advance the STT for lecturers, course administrators, and students.

CONCLUSION AND PERSPECTIVES

In this study, the STT of a small sample of 130 first-year Life Science students was investigated. The sample was from a public university in South Africa. Quantitative data were collected from the ARQ questionnaire in week one (from five factors: *achievement motivation orientation*, *learning-efficacy*, *goal orientation*, *integration and support*, and *reading behavior*), small-scale assessments in week three (*class test 1*, *tutorial 1*, and *practical 1*) and a large-scale assessment in week seven (*semester test 1*). Predictive models (*multiple linear regression*, *logistic regression*, and *decision trees*) were constructed using standard statistical practices. Through an MCCV method, the performance of the executed predictive models was assessed. The models were compared at three different time intervals (weeks one, three, and seven) to balance the trade-off between early prediction and accuracy. The balance between the time factor and accuracy is a novel aspect of this study. Also, with this study, we contribute to the discussion in the literature by offering new ideas and methods to advance the STT. Further, aligned with the call in the special issue in Educational Studies in Mathematics (Di Martino et al., 2023), this study provides unique insights from an under-represented context with a sample of teacher education students. Considering the devastating effect of COVID-19 on education, particularly in developing countries, the idea of supporting lecturers and students with data generated through predictive models in the STT seems promising.

In week one, only the decision tree model could be used to predict students who are at risk of failing the course. The prediction accuracy is low, albeit early, which provides valuable time for lecturers to design targeted support for students and for students to benefit from the support. Similar results were reported in week three.

In week seven, all predictive models could be applied to the data structure and results could be obtained. Both regression models showed the most promising results, with the multiple linear regression model outperforming the logistic regression model. On previous occasions, the logistic regression model was proven to be a powerful statistical forecast model and has achieved reliable forecasts in studies (e.g., Marbouti et al., 2016; Van Appel & Durandt, 2019). It is evident from the investigations that we could increase the prediction accuracy as we progressed through the semester. Higher accuracy values were reported in week seven compared to weeks one and three, but at a later stage in the academic semester. Thus, lecturers will have less time to take action to support at-risk students, and students will have less time to benefit from the support.

We realize the sample of life science students is small, and insignificant compared to other courses in STEM degrees with high failure rates (e.g., first-year mathematics courses). For a more comprehensive study, the models should be repeated and tested on larger samples and in other courses. An interesting aspect would also be to test the models in different contexts (e.g., samples from developing countries versus samples from developed countries).

Further ideas following from this study are to design and implement targeted intervention (or support) programs based on early prediction results, and to monitor students' progress as they are exposed to the intervention. The notions 'to predict', 'to support', and 'to monitor' at-risk students are important for role players in tertiary education (such as university administration, university management, lecturers, and students) and one notion informs the others. One reason to support this idea, from a South African perspective, is that many students cannot financially afford to repeat courses. Hopefully, role players can use the data generated from predictive models to advance the STT in science, and, more broadly, STEM courses.

Author contributions: VvA, EP, & RD: conceptualization & writing; **EP & RD:** design and methodology; **VvA:** statistical analysis. All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that ethical clearance was obtained in 2022 from the University of Johannesburg Ethics Committee. Standard measures were taken according to the literature and the university's ethical procedures. Written informed consents were obtained from the participants.

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

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Bitzer, E. (2003). Assessing students' changing perceptions of higher education. *South African Journal of Higher Education, 17*, 164-177. https://doi.org/10.4314/sajhe.v17i3.25417

Blum, W., & Leiβ, D. (2007). How do students and teachers deal with modelling problems? In C. Haines, P. Galbraith, W. Blum, & S Khan (Eds.), *Mathematical modelling: Education, engineering and economics* (pp. 222-231). Harwood. https://doi.org/10.1533/9780857099419.5.221

Braatvedt, G., & Durandt, R. (2021). A glimpse into the mathematical readiness of students undertaking a university third-year mathematics major. In M. Qhobela, M. M. E. Ntsohi, & L. G. Mohafa (Eds.), *Proceedings of the 29th Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 2-15).

Chirinda, B., Ndlovu, M., & Spangenberg, E. (2021). Teaching mathematics during the COVID-19 lockdown in a context of historical disadvantage. *Education Sciences, 11*(4), Article 177. https://doi.org/10. 3390/educsci11040177

Di Martino, P., & Gregorio, F. (2019). The mathematical crisis in secondary-tertiary transition. *International Journal of Science and Mathematics Education, 17*(4), 825-843. https://doi.org/10.1007/s10763-018-9894-y

Di Martino, P., Gregorio, F., & Iannone, P. (2023). The transition from school to university in mathematics education research: New trends and ideas from a systematic literature review. *Educational Studies in Mathematics, 113*(1), 7-34. https://doi.org/10.1007/s10649-022-10194-w

Du Plessis, A., Müller, H., & Prinsloo, P. (2005). Determining the profile of the successful first-year accounting student. *South African Journal of Higher Education, 19*(4), 684-698. https://doi.org/10.4314/sajhe.v19i4.25656

Durandt, R., Blum, W., & Lindl, A. (2021). Exploring first-year engineering student's prior knowledge in mathematics. In M. Qhobela, M. M. E. Ntsohi, & L. G. Mohafa (Eds.), *Proceedings of the 29th Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 28-31).

Efron, B., & Tibshirani, R. (1994). *An introduction to the bootstrap*. Chapman and Hall/CRC. https://doi.org/10.1201/9780429246593

Engelbrecht, J., Borba, M. C., & Kaiser, G. (2023). Will we ever teach mathematics again in the way we used to before the pandemic? *ZDM–Mathematics Education, 55*, 1-16. https://doi.org/10.1007/s11858-022-01460-5

Gale, T., & Parker, S. (2014). Navigating change: A typology of student transition in higher education. *Studies in Higher Education, 39*, 734-753. https://doi.org/10.1080/03075079.2012.721351

Geisler, S. (2021). Early dropout from university mathematics: The role of students' attitudes towards mathematics. In M. Inprasitha, N. Changsri, & N. Boonsena (Eds.), *Proceedings of the 44th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 281-288). PME.

Geisler, S., & Rolka, K. (2021). "That wasn't the math I wanted to do!"–Students' beliefs during the transition from school to university mathematics. *International Journal of Science and Mathematics Education, 19*, 599-618. https://doi.org/10.1007/s10763-020-10072-y

Glaesser, D., Holl, C., Malinka, J., McCullagh, L., Meissner, L., Harth, N. S., Machunsky, M., & Mitte, K. (2024). Examining the association between social context and disengagement: Individual and classroom factors in two samples of at-risk students. *Social Psychology of Education, 27*(1), 115-150. https://doi.org/10.1007/s11218-023-09829-4

Granger, C. W. J., & Jeon, Y. (2007). Long-term forecasting and evaluation, *International Journal of Forecasting, (23)*4, 539-551. https://doi.org/10.1016/j.ijforecast.2007.07.002

Greefrath, G., Koepf, W., & Neugebauer, C. (2017). Is there a link between preparatory course attendance and academic success? A case study of degree programmes in electrical engineering and computer science. *International Journal of Research in Undergraduate Mathematics Education, 3*, 143-167. https://doi.org/10.1007/s40753-016-0047-9

Hernandez-Martinez, P., & Williams, J. (2013). Against the odds: Resilience in mathematics students in transition. *British Educational Research Journal, 39*(1), 45-59. https://doi.org/10.1080/01411926.2011.623153

Horton, J. (2015). Identifying at-risk factors that affect college student success. *International Journal of Process Education, 7*, 83-102.

Jacobs, M., & Pretorius, E. (2016). First-year seminar intervention: Enhancing first year mathematics performance at the University of Johannesburg. *Journal of Student Affairs in Africa, 4*(1), 77-86. https://doi.org/10.14426/jsaa.v4i1.146

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning*. Springer. https://doi.org/10.1007/978-1-0716-1418-1

Jansen, E. P. W. A., & Van der Meer, J., (2012). Ready for university? A cross-national study of students' perceived preparedness for university. *Australian Educational Researcher, 39*, 1-16. https://doi.org/10.1007/s13384-011-0044-6

Kuhn, M., & Johnson, K. (2016). *Applied predictive modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Lemmens, J. (2010). *Students' readiness for university education* [PhD thesis, University of Pretoria].

Leong, E., Mercer, A., Danczak, S. M., Kyne, S., & Thompson, C. D. (2021). The transition to first year chemistry: Student, secondary and tertiary educator's perceptions of student preparedness. *Chemistry Education Research and Practice, 22*, 923-947. https://doi.org/10.1039/D1RP00068C

Lim, W., Bae, Y., & Know, O. (2022). The development of sociomathematical norms in the transition to tertiary exam-oriented individualistic mathematics education in an East Asian context. *Educational Studies in Mathematics, 113*, 57-78 https://doi.org/10.1007/s10649-022-10203-y

Marbouti, F., Diefes-Dux, H. A., & Madhavan, K. (2016). Models for early prediction of at-risk students in a course using standards-based grading. *Computers and Education, 103*, 1-15. https://doi.org/10.1016/j.compedu.2016.09.005

Meyer, W. J. (2012). *Concepts of mathematical modeling*. Dover Publications, INC.

Morton, W., & Durandt, R. (2023). Learning first-year mathematics fully online: Were students prepared, how did they respond? *EURASIA Journal of Mathematics, Science and Technology Education, 19*(6), Article em2272. https://doi.org/10.29333/ejmste/13189

Mullis, I. V. S, Martin, M. O., Foy, P., Kelly, D. L., & Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science*. TIMSS & PIRLS International Study Center.

Nicholson, N. (1990). The transition cycle: Causes, outcomes, processes and forms. In S. Fisher, & L. Cooper (Eds.), *On the move: The psychological effects of change and transition* (pp. 83-108). John Wiley.

Niss, M., & Blum, W. (2020). *The learning and teaching of mathematical modelling*. Routledge. https://doi.org/10.4324/9781315189314

Nouwen, W., & Clycq, N. (2020). Assessing the added value of the self-system model of motivational development in explaining school engagement among students at risk of early leaving from education and training. *European Journal of Psychology of Education, 36*(2), 243-261. https://doi.org/10.1007/s10212-020-00476-3

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33-40. https://doi.org/10.1037/0022-0663.82.1.33

Ribeiro, L., Rosário, P., Núñez, J. C., Gaeta, M., & Fuentes, S. (2019). First-year students background and academic achievement: The mediating role of student engagement. *Frontiers in Psychology, 10*, Article 2669. https://doi.org/10.3389/fpsyg.2019.02669

Rossato, L., Morotti, A. C. V., & Scorsolini-Comin, F. (2024). Transition and adaptation to higher education in Brazilian first-year nursing students. *Journal of Latinos and Education, 23*(1), 2-12. https://doi.org/10.1080/15348431.2022.2102499

Shan, G. (2022). Monte Carlo cross validation for a study with binary outcome and limited sample size. *BMC Medical Informatics and Decision Making, 22*(270), 1-15. https://doi.org/10.1186/s12911-022-02016-z

Sun, J. C.-Y., Joo Oh, Y., Seli, H., & Jung, M. (2017). Learning behavior and motivation of at-risk college students: The case of a self-regulatory learning class. *The Journal of At-Risk Issues, 20*(2), 12-24.

Tall, D. (2008). The transition to formal thinking in mathematics. *Mathematics Education Research Journal, 20*(2), 5-24. https://doi.org/10.1007/BF03217474

Van Appel, V., & Durandt, R. (2019). Investigating possibilities of predictive mathematical models to identify at-risk students in the South African higher education context. *Perspectives in Education, 37*(2), 1-15. https://doi.org/10.18820/2519593X/pie.v37i2.1

van der Put, C. E. (2020). The development of a risk and needs assessment instrument for truancy. *Children and Youth Services Review, 109*, Article 104721. https://doi.org/10.1016/j.childyouth.2019.104721