The Version of Record of this manuscript has been published and is available in *Studies in Higher Education*, 2024
https://www.tandfonline.com/doi/full/10.1080/03075079.2024.2338262

Predicting performance in exams and deep approach to learning in first year university students: A new look at academic success

Manuel Bächtold[a,b], Jacqueline Papet[a,b], Dominique Barbe Asensio[a], André Mas[a,c], Sandra Borne[a], & Appolinaire Ngoua Ondo[a]

[a]*OTP, SCSIP, University of Montpellier, France*
[b]*LIRDEF, University of Montpellier & University Paul Valéry Montpellier, France*
[c]*IMAG, University of Montpellier & CNRS, France*

Abstract
This study calls for a broadening of the perspective on academic success. While passing exams is an essential objective of higher education, it should not overshadow another important objective which is the development of students' skills, such as becoming curious, autonomous and reflective in the learning process. This study used Academic Performance in Exams (APE) and Deep Approach to Learning (DAL) as measures related to these two objectives. The aim was to identify and compare the factors that may influence APE and DAL. The study was conducted on first-year students (2011) at a French university. It was based on a random forest algorithm and took into account a wide range of factors belonging to different dimensions: demographics, social background, educational background, context of the educational programme, behavioural engagement, social environment, psychological and cognitive characteristics. The results show that the most important factors in predicting APE are the educational programme undertaken, student's educational background and parents' occupation. DAL was not found to be an important factor in APE. Regarding the prediction of DAL, the results point to the predominant weight of intrinsic motivation and the important weight of elaborated epistemic beliefs. In contrast, demographics and behavioural engagement were found to have negligible weight in predicting both APE and DAL. These findings raise questions about the type of success that is valued in the first year of university and call for reflection on assessment methods. They also allow the identification of levers that teachers can activate to support first year students.

Keywords: approach to learning; motivation; epistemic beliefs; difficulties; social support.

1. Introduction

Academic success in first year of university has become a major issue for higher education institutions and policies (Clerici et al. 2015). First-year students run a greater risk of dropping out or failing to achieve the grades required to reach the next year (Fokkens-Bruinsma et al. 2021). In addition to choosing the right educational programme for them, first-year students face a challenging transition from high school to higher education. This transition involves changes in their educational environment, such as new learning tasks, new relations with the teachers, new social networks, new time management, and new engagement in studies (De Clercq et al. 2017). This transition also implies a new way of thinking (Entwistle 2009), of understanding, experiencing, and conceptualizing the world (Ramsden 2003). Students must undergo an intellectual, methodological and sociological metamorphosis in order to succeed in their first year (Paivandi 2015).

Recently, some researchers (van der Zanden et al. 2019) have taken a new look at this issue of academic success, by emphasizing that it should not be limited to Academic Performances in Exam (APE). University education has a dual challenge: to help students achieve an academic degree, but also develop the psychological and cognitive skills that are essential for their future life as citizens and for their professional development. This second challenge implies students becoming curious, autonomous and reflective in their learning. In this respect, encouraging students to adopt a Deep Approach to Learning (DAL) (Biggs and Tang 2011), that contributes to the development of these skills, can be seen as an essential aim of university education. According to this shift in perspective, DAL should not only be considered as a possible factor of academic success, measured in terms of exam grades, but as a complementary outcome of higher education that deserves to be examined for its own sake.

Numerous studies have been carried out to determine the influence of a range of variables on APE. In a systematic literature review, Hellas et al. (2018) distinguished various kinds of factors in APE: demographics, family background, educational background, course data, working conditions, student motivation, and psychological, affective and learning scales. Several studies have compared the relative weight of some of these variables, most of them based on linear regressions (e.g., Clerici et al. 2015; Pinxten et al. 2015; Sothan 2019). However, all the variables have not yet been taken into account together in a same study for a systematic comparison. Multiple studies have also been conducted to determine the factors that favour DAL, taking into account both student factors (such as age, gender, motivation or self-esteem) and contextual factors related to the teaching environment (Baeten et al. 2010). Concerning this output variable, fewer studies have compared the relative weights of the different factors.

The aim of this study is to determine in a more systematic way the relative weight of the factors that may influence APE, by integrating a wide range of potential factors, identified in the literature, into a single predictive model. This model is based on a Random Forest algorithm (RF), which has a twofold advantage: it does not presuppose a linear relationship between the input variables and the output variable, and it can include both quantitative and nominal variables. According tothe change in perspective mentioned above, this study also aims to determine the relative weights of the variables that may influence DAL, with an equivalent predictive model. The two models can thus be compared and contribute to a discussion of the links between APE and DAL.

2. Factors identified in previous research

Different types of factors have been identified in the literature that may favour APE and/or DAL. They can be classified into the following categories: demographics, social background, educational background, context of the educational programme, social environment, behavioural engagement, psychological and cognitive characteristics.

2.1 Demographics

Either positive or negative correlations were found between age and APE (Cassidy 2012; Craft 2019), gender and APE (Bruinsma 2004; Clerici et al. 2015), and gender and DAL (Berberoglu and Hei 2003; Mattick et al. 2004). The correlation found between age and DAL are often positive (Gijbels et al. 2005), but sometimes not significant (Duff et al. 2004). A relation was

found between the country of origin and APE, with native students having an advantage over foreign students (Clerici et al. 2015; Craft 2019).

2.2 Social background

The socio-economic status (SES) of students' families has been found to be positively correlated with APE in several studies (Brinbaum et al. 2018; Pinxten et al. 2015; Sothan 2019), with one counter-example of a non-significant relationship (Craft 2019). Regarding the relationship between SES and DAL, one study found a negative correlation (Suphi and Yaratan 2012), while another found no significant correlation (Schrempft et al. 2021).

2.3 Educational background

Several studies have found that high school grades positively predict APE at university (Clerici et al. 2015; Sothan 2019). Little research has looked at the relationship between high school grades and DAL at university, but there is one study that found a negative relationship (Duff et al. 2004).

2.4 Context of the educational programme

The field of study has been found to have an influence both on APE (Clerici et al. 2015; Fokkens-Bruinsma et al. 2021) and DAL (Nelson Laird et al. 2008). DAL is more prevalent in soft fields than in hard fields. The relationship between the field and APE is more complex.

2.5 Behavioural engagement

There are mixed results regarding the relationship between effort (i.e. time spent studying) and APE, with several studies finding a positive relationship (Diseth et al. 2010; Dupont et al. 2015; Sothan 2019) and others finding a negative relationship (Fokkens-Bruinsma et al. 2021). Positive correlations were found between effort and DAL (Diseth et al. 2010; Román et al. 2008).

2.6 Social environment

The influence of the social environment can be measured in terms of how students perceive it. Perceived social support was found to have a significant effect on APE (DeBerard et al. 2004). More specifically, supervisors, family and institutional social support, but not peer social support, were found to be positively correlated with APE (Dupont et al. 2015). Another study found that family support was positively correlated with DAL, but not with APE (Román et al. 2008). Perceived social climate, which encompasses atmosphere, culture, values, organizational, instructional and interpersonal dimensions, was found to correlate with APE but only in some classroom contexts (Rania et al. 2014).

2.7 Psychological and cognitive characteristics

Intrinsic and more broadly self-determined motivation (SDT) (Ryan and Deci 2000, 2020) were found to be positive predictors in APE (Fokkens-Bruinsma et al. 2021). Many studies have found a positive correlation between intrinsic motivation and DAL (Chue and Nie 2016; Liu et al. 2015; Minbashian et al. 2004), which is why both variables have often been combined into a single construct (which will be avoided below for conceptual clarity).

Self-efficacy has also been found to correlate positively with both APE (Greco et al. 2022; Guo et al. 2022; Román et al. 2008) and DAL (Lin and Tsai 2013; Liu et al. 2015; Román et al. 2008). An alternative way of looking at how students perceive their abilities when studying is to consider their perceived difficulties (Ainscough et al. 2018; Cameron and Rideout 2020; Trautwein and Bosse 2017). However, to our knowledge, the relationship between these perceived difficulties and APE or DAL has not yet been investigated.

Evaluativist epistemic beliefs (evaluativism), which are the most elaborated beliefs on the nature of knowledge and the processes of knowing (Hofer and Pintrich 1997), have been found to correlate positively with APE for some kinds of beliefs (e.g., uncertainty of knowledge), but not for others (e.g., subjectivity of knowledge) (Aditomo 2018; Guo et al. 2022; Lonka et al. 2021). A more straightforward positive relationship has been found between evaluativism and DAL (Chiu et al. 2016; Lehmann 2022; Lin et al. 2012).

Finally, DAL can also be considered as a factor in APE. Several studies found a positive correlation between the two variables (Diseth et al. 2010; Liu et al. 2015; Schrempft et al. 2021), while some studies found a negative correlation (Bruinsma 2004) or no significant correlation (Gijbels et al. 2005; Minbashian et al., 2004).

3 How to compare the relative weight of the factors?

Given the large number of factors influencing APE and DAL identified in the research, the question arises as to their respective weight. In the case of APE, but not DAL, several studies have been carried out to compare these factors. However, each of these studies is restricted to a limited number of the above factors. Moreover, most of them are based on linear regressions (Cassidy 2012; Duff et al. 2004; Sothan 2019), which in some studies are used to develop structural equation models (Diseth et al. 2010; Dupont et al. 2015; Román et al. 2008). A problem is that the relationships between factors and APE are not necessarily linear (Diseth 2002; Mouratidis et al. 2021; Musso et al. 2020), which may also be the case for DAL.

More recently, researchers have used machine learning algorithms, such as RF, Classification Tree, or Neural Network, to predict APE and address these two limitations (Beaulac and Rosenthal 2019; Cannistra et al. 2022; Musso et al. 2020). These algorithms allow to include a large number of factors, quantitative or nominal, in the same model and without assuming linear relations with the target variable. Moreover, they enable to compare the predictive weight of the factors in the model, and to reach a high predictive power.

4. The present study

Focusing on the first year of university, the present study considers two objectives of university education: helping students succeed in their exams, and developing psychological and cognitive skills necessary for their future life as citizens and their professional development. These two objectives can be related to two different outcome variables: APE and DAL. In order to predict these two variables, we propose to use the RF algorithm and to consider a large set of factors that fall under the different categories identified in the literature: demographics, social background, educational background, context of the educational programme, behavioural engagement, social environment, psychological and cognitive characteristics. Considering these sets of factors jointly in the same models provides an opportunity to compare their

respective influence on APE and DAL in a more systematic way than in previous research. Accordingly, this study aims to address the following two research questions:

- RQ1: What is the relative weight of the different types of factors in predicting APE for first year university students?
- RQ2: What is the relative weight of the different types of factors in predicting DAL for first year university students?

A more specific question related to RQ1 is whether or not DAL itself is an important factor in APE. By investing RQ1 and RQ2 in the same study, we will also be able to compare the weight of the factor predicting APE and DAL, and determine whether there are some common important factors and, accordingly, common possible levers to activate.

5. Method

5.1 Participants and procedure

A total of 7301 first-year students at a French university were invited to complete a questionnaire and 2011 students responded to all questions (27.5% of respondents). This sample was composed of students enrolled in ten different educational programmes, in science and/or humanities. The participants had a mean age of 18.35 years (SD = 1.46), and 58.8% were female.

The questionnaire was administered electronically two months after the start of their first year of study. This delay had two purposes: to avoid including in the sample students who change programs during the period that authorizes it administratively, and to include students who drop out during the year (the later in the year the questionnaire being administered, the fewer students who drop out being taken into account). Prior to completing the questionnaire, all participants provided informed consent. Students with missing data and problematic outliers were removed from the sample.

5.2. Data collection

The study is based on data provided by the university administration and data collected through the questionnaire. The latter was composed of 61 closed-ended questions. Some of them were adapted from the literature, while others were created (for the complete questionnaire, see Supplemental material A). The question wording was adjusted after a qualitative pre-test based on interviews with five students and three successive tests with 980 students in total.

5.3 Outcome variables

To run the RF algorithm, the outcome variables have been defined in binary form. The first outcome variable was APE, defined as passing or not passing midterm exams. According to university regulations, a student passes his or her exams if the average of his or her grades is equal to or higher than 10 on a scale of 0 to 20. The data regarding students' grades were supplied by the university administration. It should be noted that the midterm exams are the first exams that students take and that the exams they take at the end of the first year are generally similar in format.

The second outcome variable was DAL, defined as above or below the mean of the students on a scale based on thirteen items (McDonald's ω = .87). The DAL scale consisted of four subscales: integrative approach expressing a structured thinking and composed of five

items (ω = .80), reflective approach composed of four items (ω = .69), interactions-with-peers approach composed of two items (ω = .86), and interactions-with-teachers approach composed of two items (ω = .80). Integrative and reflective approaches included items adapted from the Biggs et al. (2001) and Entwistle and Mac Cune (2013) questionnaires. The other two subscales are new and describe students' intent to better understand what is being taught through social interactions. For these items related to DAL, students were asked to respond on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

5.4 Factor variables

5.4.1 Demographics, social background, educational background, and context of the educational programme

The university administration supplied demographics on the students (age, gender, native or foreigner), on their social background (parents' occupation, whether they had a scholarship, or a job), on their educational background (high school graduation mark, type of high school diploma, whether they were just graduated from high school, or repeated their first year of university), and on the context of the educational programme (educational programme undertaken, type of degree, and whether their educational programme corresponded to their first wish).

5.4.2 Behavioural engagement

Two items of the questionnaire provided data on student behavioural engagement: course attendance (7-point Likert scale measuring frequency), and the work at home for studies (with five time intervals).

5.4.3 Social environment

The social environment consisted of the perceived social support and perceived social climate. The perceived social support scale was constructed as the mean size of the network (i.e. number of different sources) for five types of support (financial, learning, course guidance, confidence, and project) (ω = .80). Regarding the perceived social climate of the educational programme, students had to position a cursor on a scale ranging from 1 meaning stressful to 7 meaning reassuring, and on another scale ranging from 1 meaning free to 7 meaning supervised.

5.4.4 Psychological and cognitive characteristics

Motivation was measured by means of the Vallerand et al. (1989) scale, in line with SDT and adapted to the university context. It was based on seven subscales, each composed of four items: intrinsic motivation to know (ω = .88), to stimulation (ω = .83), to accomplishment (ω = .87), identified extrinsic motivation (ω = .80), introjected extrinsic motivation (ω = .83), external extrinsic motivation (ω = .80), and amotivation (ω = .86).

Evaluativism consisted of four items (ω = .62) corresponding to the four kinds of epistemic beliefs distinguished by Hofer and Pintrich (1997): beliefs concerning the uncertainty, complexity, source, and justification of knowledge. Note that such a relatively low value for the reliability coefficient is not unusual (DeBacker 2008). In several studies, the values found and used for measures of epistemic beliefs are around 0.6 or even below 0.6 (e.g. supplementary materials in Schiefer et al. 2022). This can be explained by the fact that, for some individuals, epistemic beliefs may not form a fully coherent system (Schommer 1990).

Perceived difficulties in studying consisted of seven items (ω = .83) that reflect the main difficulties identified in the literature (Ainscough et al. 2018; Cameron and Rideout 2020; Trautwein and Bosse 2017): difficulties with course guidance, in understanding courses, in understanding course issues, in understanding course instructions, due to content complexity, with learning methods, and in organizing work out of class.

For all the items related to motivation, amotivation, evaluativism, and perceived difficulties, students were asked to respond on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).

For APE as the outcome variable, the four subscales of DAL mentioned above were also considered.

5.5 The two models

As two outcome variables were investigated in the study, two different models were developed: a model with 31 factors for APE (Figure 1) and another with 27 factors for DAL (Figure 2).

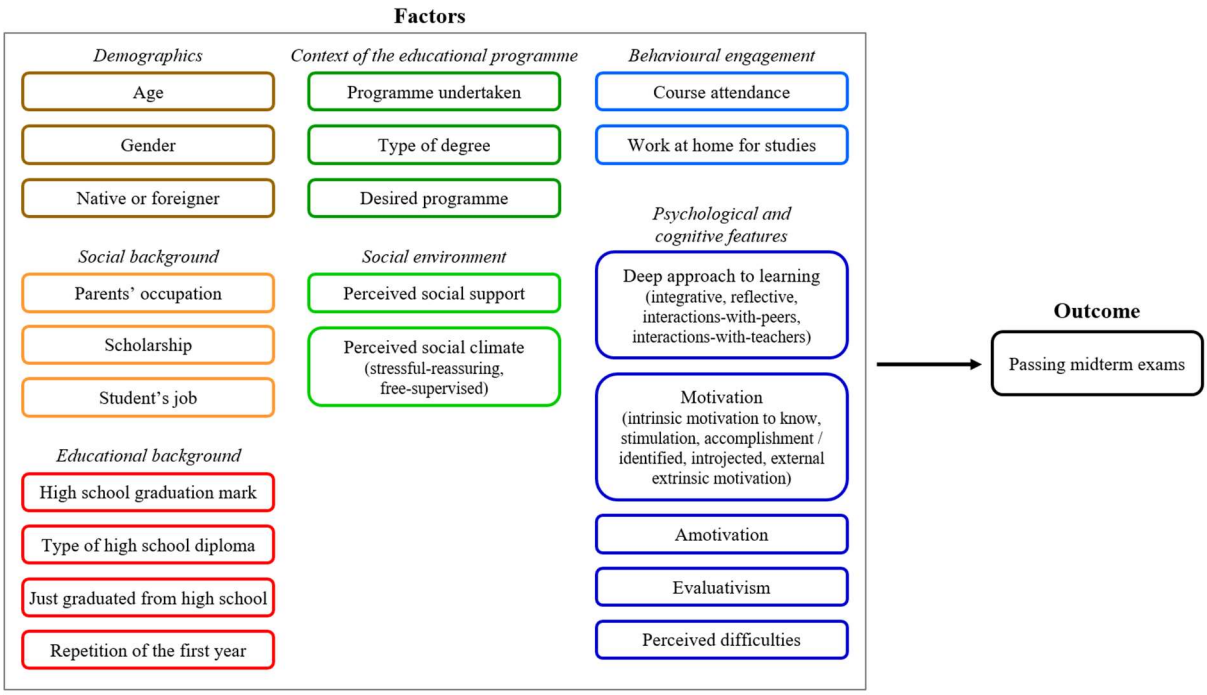

Figure 1. The factors considered to predict APE.

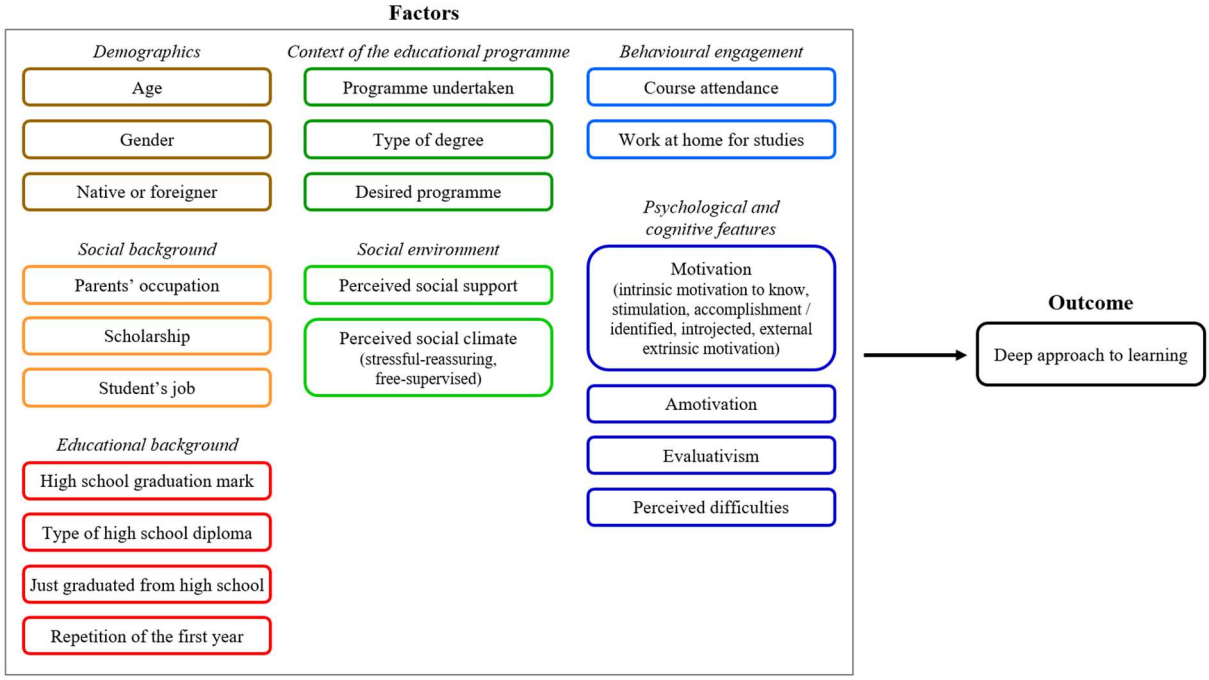

Figure 2. The factors considered to predict DAL.

5.6 Statistical analysis

To investigate the relationships between the factor variables and the outcomes variables, we used the RF algorithm, that involves a training phase with a randomly selected part of the observations, the bag sample, and a prediction phase with the remaining observations, the out-of-bag sample (Breiman 2001; Rosenbusch 2021). It is based on the method of binary decision trees, which consists of dividing the set of observations thanks to successive binary questions. Each question relates to a factor variable. The order of the questions is determined by their discriminatory power relative to the output variable, with the first question being the most discriminatory. The successive questions allow the construction of subsets of observations that are increasingly homogeneous with respect to the output variable (for example, subsets of students who, for the most part, passed their midterm exams and subsets of students who, for the most part, did not passed them). RF combines a large number of such decision trees, each of which covers a sample of randomly selected observations of the bag. For each split in a given tree, a subset of input variables is randomly selected. To test the RF model, observations from the out-of-bag sample are used. For each of these observations, the model makes a prediction about the output variable by taking into account the predictions of all decision trees and retaining the most frequent prediction. A level of prediction accuracy can then be associated to the model. The contribution of each variable to the prediction can be determined using the mean decrease of the Gini Index, which indicates the purity of a dataset's partition. This value can be translated into a percentage contribution to the prediction of the outcome variable in the model. RF allows for the inclusion of a large number of variables, which may be quantitative (continuous or discrete) or nominal, in a single model and does not make the assumption of linear relationships with the outcome variable.

The RF analyses were performed with R software and the randomForest package. To optimize the prediction accuracy, the following three parameters were tuned for each RF model: number of randomly selected input variables for each split, number of decision trees, and number of splits. The reliability, descriptive and correlation analyses were performed by means of JASP software (Version 0.16.2).

6. Results

Descriptive statistics for all variables in both models, as well as correlations between the two output variables, can be found in Supplemental material B.

6.1. Relative weight of the factors predicting APE

In the case of APE, the RF analysis resulted in a model with a prediction accuracy of 74.9%. In this model, the percentage of contribution of each variable to the prediction of APE is given in Figure 3 (for the percentages and corresponding means of the Gini index, see Supplemental material C).

According to this model, the most important factor is the educational programme undertaken (contributing 13.4% to the prediction in the model). Comparing the passing rate between the educational programmes considered in the study, the rate is highest in the case of a preparatory degree for teaching and a Technology University Diploma, and lowest in the case of the Bachelors of economics and science. The second most important factor is the high school graduation mark (9.3%). The relationship between this factor and APE is positive, as indicated by the value of their correlation (Spearman's $\rho = 0.361$, $p < .001$). A third important factor is the occupation of the parents (8.6%). The importance of this factor can be illustrated by the following examples: the percentage of students who succeed in midterm exams is 71.0% when both parents are executives or intermediaries, 60.6% when both parents are employees or workers, and 46.2% when both parents are unemployed. Furthermore, according to this model, DAL is a secondary factor in APE (2.75% to 3.5% depending on the subscale considered). This means that the exams reward students who adopt a DAL only to a limited extent. Demographic data (age, gender and country of origin) are among the least decisive factors. The same applies to behavioural engagement.

6.2. Relative weight of the factors predicting DAL

In the case of DAL, the RF analysis resulted in a model with a prediction accuracy of 77.0%. In this model, the percentage of contribution of each variable to the prediction of DAL is given in Figure 4 (for the detailed percentages and corresponding means of the Gini index, see Supplemental material C).

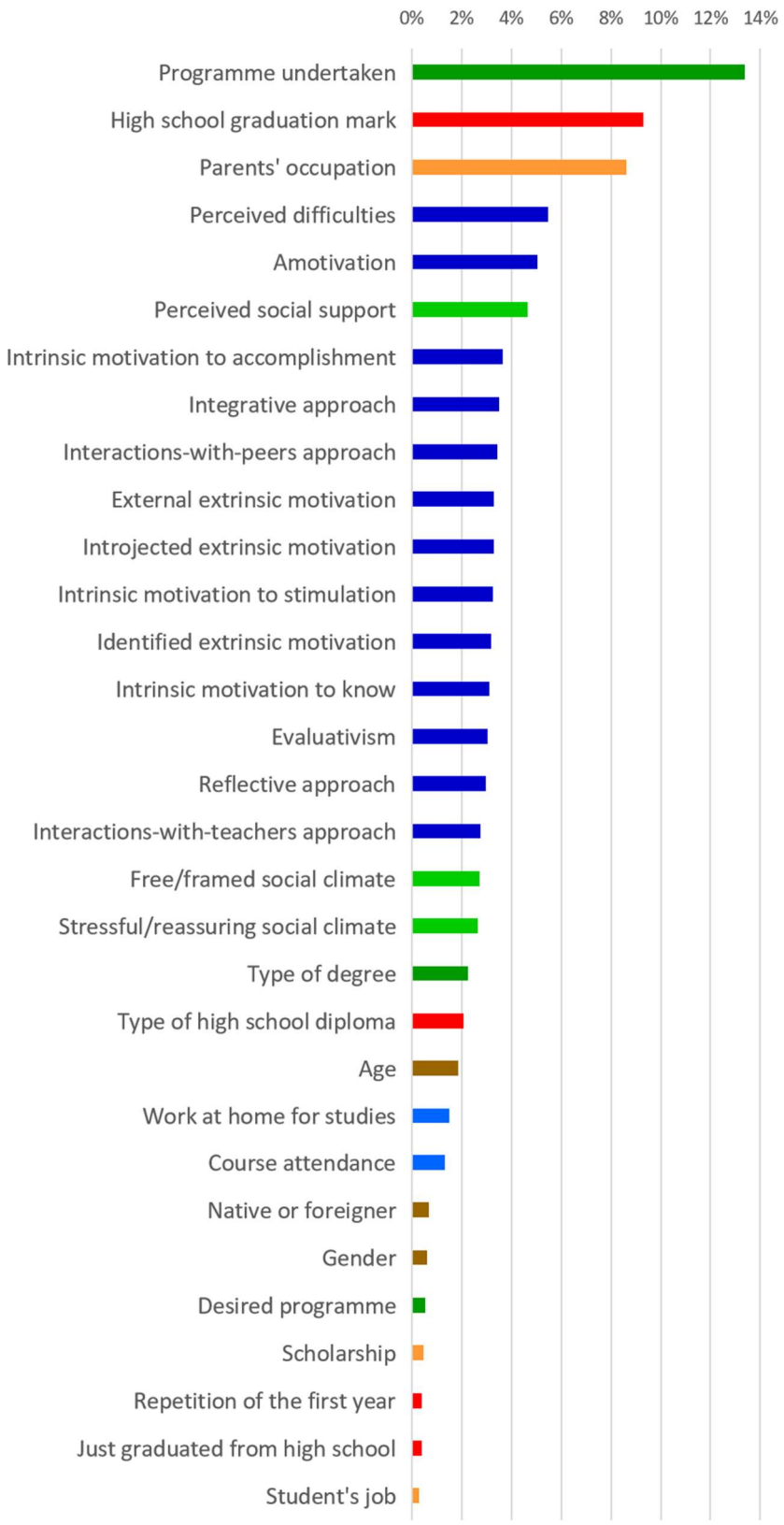

Figure 3. Contribution of each variable to the prediction of APE (passing midterm exams).

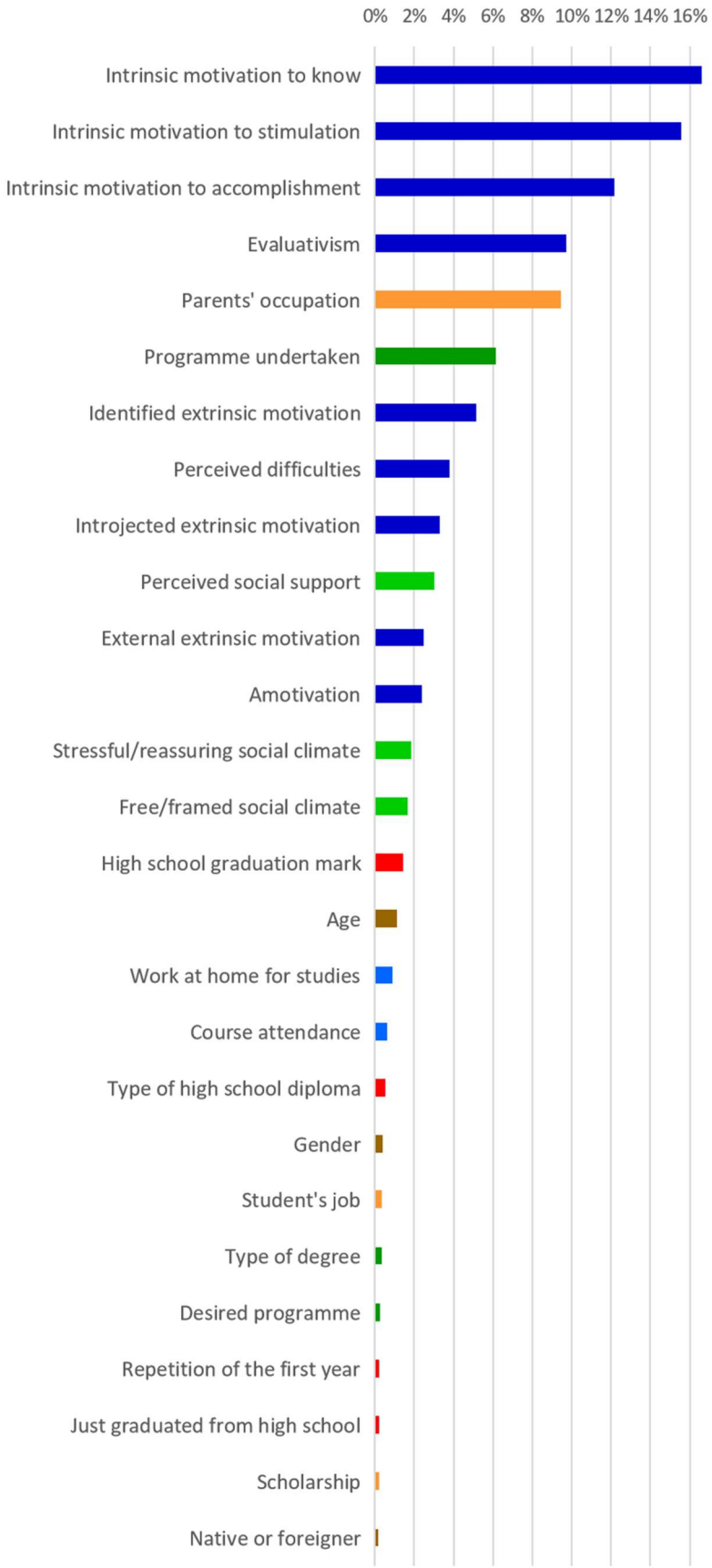

Figure 4. Contribution of each variable to the prediction of DAL.

In this second model, the most important factors are the three forms of intrinsic motivation: to know (16.6%), to stimulation (15.6%), and to accomplishment (12.2%). Evaluativism is also an important factor, although somewhat weaker (9.7%). The relationship between this factor and DAL is positive, as indicated by the value of their correlation (ρ = 0.375, $p < .001$). Another factor with roughly the same weight as evaluativism is the parents' occupation (9.4%). The relationship between this factor and DAL appears to be inverse compared to its relationship with APE, as shown by the following examples: the percentage of students with a DAL score above the mean is 69.2% when both parents are unemployed, 52.7% when both parents are employees or workers, and 52.3% when both parents are executives or intermediaries. Finally, as with APE, demographics and behavioural engagement are among the least decisive factors.

7. Discussion

The aim of this study was to identify and compare the factors that may influence APE and DAL in the first year of university, using the RF algorithm and taking into account a wide range of factors belonging to a variety of dimensions: demographics, social background, educational background, context of the educational programme, behavioural engagement, social environment, psychological and cognitive characteristics.

Regarding APE, the results show the predominant importance of the educational programme undertaken in passing the exam. Behind this variable lie several possible factors, such as discipline, admission conditions, student-teacher ratio, promotion size and teaching methods. A controlled study would be needed to determine which of these factors are the most decisive. It should be noted, however, that the two educational programmes considered in the study with the highest exam success rates are characterized by a selection process based on students' academic records on entry, a high student-teacher ratio (i.e. number of teachers per student) and a small class size, which is not the case for the two educational programmes with the lowest success rates.

The findings also show the important weight of students' educational background and confirm previous studies (Clerici et al. 2015; Sothan 2019). Students who perform better in high school grades tend also to perform better in exams at university. This means that high performance in high school exams may be a factor of stability for students in coping with the difficult transition to the university system.

Parental occupation is a third important factor in predicting APE. The higher the socio-economic level of the parents' occupations, the more likely students are to succeed in their exams. This result is in line with those of several studies and in particular with that of a study conducted in the same country, France (Brinbaum et al. 2018). This confirms that, in this country, educational inequalities, which are strongly linked to students' socio-economic backgrounds, extend into the first year of university.

According to our study, DAL is not a major factor in APE. This finding raises questions about the type of success that is valued in the first year of university. It calls for reflection on assessment methods. Indeed, according to several studies, assessments which require the memorisation of knowledge rather than conceptual mastery or detailed answers rather than relating content are not beneficial to students who adopt DAL and, consequently, do not

encourage students to adopt such an approach to learning during their studies (Bruinsma 2004; Dahlgren et al. 2009; Minbashian et al. 2004).

With respect to DAL, the results point to the preponderant weight of the three forms of intrinsic motivation (i.e. to knowledge, stimulation and achievement), in line with the findings of previous studies (Chue and Nie 2016; Liu et al. 2015; Minbashian et al. 2004). These results point to possible levers that teachers can activate to lead their students to adopt a more in-depth approach to learning: these levers would consist not only in arousing their interest in the content taught, but also in offering them activities that are stimulating and that give them a sense of achievement. Furthermore, we can see that the decreasing weight of the three forms of extrinsic motivation (i.e. identified, introjected and external) echoes their decreasing value on the SDT self-determination scale (Ryan and Deci 2000). In other words, the less students' motivation expresses a form of autonomy in the regulation of their activities, the less likely it is to favour DAL.

The findings also highlight the important weight of evaluativism, i.e. elaborated epistemic beliefs. The relatively strong link between evaluativism and DAL has been established by several previous studies (Chiu et al., 2016; Lehmann, 2022; Lin et al., 2012). This result points to a second lever for teachers: this would consist in offering students opportunities to discuss the nature of the knowledge taught and the process of constructing and validating this knowledge. Indeed, a number of studies have shown that epistemic beliefs in science remain poor or naïve if the various aspects of the nature of scientific knowledge and its construction and validation are not explicitly addressed (Khishfe 2023).

The results show the relatively high predictive weight of parents' occupation for DAL. However, this factor plays a different role for DAL than for APE. While the relationship between the family's socio-economic level is positive with APE, it is negative with DAL. This result may seem surprising, but it is consistent with the results found in two studies (Schrempft et al. 2021; Suphi and Yaratan, 2012). One possible explanation could be that socio-economically disadvantaged students are more inclined to learn in depth in order to compensate, consciously or not, for their lower chances of academic success. This result could be related to a study (Macaulay et al., 2023) showing that students from low socioeconomic backgrounds express strong motivation to develop strategies for success.

The study also shows that some factors identified in the literature are weaker than others. This is the case in particular for perceived difficulties and perceived social support, which are relatively less important, although not negligible, in predicting APE and DAL. These results point to other possible levers for teachers, but whose potential influence in favouring APE and DAL is weaker. The first lever is to take greater account of the different types of difficulties experienced by students in their studies, such as difficulties in understanding course issues or course instructions, difficulties due to the complexity of the content, difficulties with learning methods or organising work out of class (Ainscough et al. 2018; Cameron and Rideout 2020; Trautwein and Bosse 2017). The second lever is to provide more support to students, be it for learning, course guidance or self-confidence (Dupont et al. 2015).

The study found that behavioural engagement, measured by course attendance and work at home for studies, has a relatively weak predictive value for both APE and DAL. In terms of predicting APE, the weight of behavioural engagement is lower than that of the different subscales of DAL. This finding suggests that the quality of a student's work is more important

than its quantity. It should be noted that this finding contrasts with other studies in which behavioural engagement is a more important factor in exam success (Diseth et al. 2010; Sothan 2019). A possible explanation for this discrepancy may lie in the dependence of the strength of this factor on the educational programme or discipline.

Finally, the findings indicate that demographics (age, gender, and country of origin) have negligible weight in predicting both APE and DAL. With regard to country of origin, this result should be treated with caution as foreign students represented only 5.62% of the total sample. As for age and gender, their weak influence is consistent with the fact that contradictory results have been found in the literature, whether in relation to APE (Bruinsma 2004; Cassidy 2012; Clerici et al. 2015; Craft 2019) or DAL (Berberoglu and Hei 2003; Duff et al. 2004; Gijbels et al. 2005; Mattick et al. 2004). If the weight of these factors is low, the nature of their relationship with APE or DAL, whether positive or negative, could be more dependent on the context.

8. Educational implications

This study calls for broadening the perspective on the issue of academic success, in line with previous authors (van der Zanden et al. 2019). While passing exams and graduation of students is a major objective of higher education, it should not overshadow another objective that is linked to an important development of students: to become curious, autonomous and reflective in the learning process. In this study, which focused on the first year of university, APE and DAL were used as measures related to these two objectives. According to the findings, these two objectives do not coincide, as DAL is not the most important factor in APE. This result might lead teachers to question their assessment procedures and possibly revise them, so that they reward students adopting DAL more and thereby encourage them to adopt DAL in their learning. In addition, this study allowed us to identify several possible levers that teachers can use to encourage their students to adopt DAL: levers consisting of increasing their intrinsic motivation, in its three forms (i.e. to know, to stimulation and to accomplishment), and the lever consisting of enriching their epistemic beliefs.

9. Limitations and future directions

This study has several limitations. The RF analyses were conducted by distinguishing students in a binary manner with respect to each outcome variable. It is possible to refine this analysis by partitioning the students into more groups and looking for the most predictive factors that discriminate between these groups. Furthermore, the RF algorithm offers a simple predictive model, linking all the factors directly to the output variable. Some relationships could be indirect, meaning that one factor could influence the output variable through the mediation of another variable. In this respect, structural equation modelling can provide complementary insights. Finally, the results obtained concern only first-year students in a French university. This last limitation points to future research directions. In particular, the present study would deserve to be replicated in first-year students at other universities to determine the extent to which the results can be generalized. It could also be replicated in the same university in the second and third years of study, in order to examine the evolution of the relative predictive weight of the different factors. In particular, we can wonder whether, in the subsequent years of university studies, the high school graduation mark remains an important factor in APE,

whether the nature of assessments is changing and tending to give greater importance to DAL, whether the link between evaluativism and DAL remains relatively strong, or whether behavioural engagement is becoming a more decisive factor in APE and DAL.

Funding
This work was supported by the Observatoire de la Transformation Pédagogique (OTP) de l'Université de Montpellier (UM).

Disclosure statement
The authors report there are no competing interests to declare.

Ethical considerations
Prior to completing the questionnaire of the study, all participants provided informed consent in compliance with the European General Data Protection Regulation (EU) 2016/679 and under the supervision of the Data Protector Officer of the university. The study received formal approval of the Research Ethics Committee of the University of Montpellier.

Data availability
In line with the commitment made to the students surveyed, the complete raw data file is not provided online. They are available on request, as are the codes used.

References
Aditomo, A. 2018. "Epistemic beliefs and academic performance across soft and hard disciplines in the first year of college." *Journal of Further and Higher Education* 42 (4): 482–496.

Ainscough, L., Stewart, E., Colthorpe, K., and K. Zimbardi. 2018. "Learning hindrances and self-regulated learning strategies reported by undergraduate students: identifying characteristics of resilient students." *Studies in higher education* 43 (12): 2194–2209.

Baeten, M., Kyndt, E., Struyven, K., and F. Dochy. 2010. "Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness." *Educational Research Review* 5 (3): 243–260.

Beaulac, C., and J. Rosenthal. 2019. "Predicting university students' academic success and major using random forests?" *Research in Higher Education* 60: 1048–1064.

Berberoglu G. and L. Hei. 2003. "A comparison of university students' approaches to learning across Taiwan and Turkey." *International Journal of Testing* 3 (2): 173–187.

Biggs, J., Kember, D., and D. Leung. 2001. "The revised two factor study process questionnaire: R-SPQ-2F." *British Journal of Educational Psychology* 71 (1): 13–149.

Biggs, J., and C. Tang. 2011. *Teaching for quality learning at university*. Open University Press.

Breiman, L. 2001. "Randon Forests." *Machine learning* 45: 5–32.

Brinbaum, Y., Hugrée, C., and T. Poullaouec. 2018. "50% to the bachelor's degree... but how? Young people from working-class families at university in France." *Economics and Statistics* 499: 79–105.

Bruinsma, M. 2004. "Motivation, cognitive processing and achievement in higher education." *Learning and Instruction* 14 (6): 549–568.

Cameron, R., and C. Rideout. 2020. "'It's been a challenge finding new ways to learn': first-year students' perceptions of adapting to learning in a university environment." *Studies in Higher Education* 47 (3): 668–682.

Cannistra, M., Masci, M., Leva, F., Agasiti, T., and A-M. Paganoni. 2022. "Early-predicting dropout of university students: an application of innovative multilevel machine learning and statistical techniques." *Studies in Higher Education* 47 (9): 1935–1956.

Cassidy, S. 2012. "Exploring individual differences as determining factors in student academic achievement in higher education." *Studies in higher education* 37 (7): 793–810.

Chiu, Y.-L., Liang, J.-C., Hou, C.-Y., and C.-C. Tsai. 2016. "Exploring the relationships between epistemic beliefs about medicine and approaches to learning medicine: a structural equation modeling analysis." *BMC Medical Education* 16: article 181.

Chue, K. and Y. Nie. 2016. "International students' motivation and learning approach: A comparison with local students." *Journal of International Students* 6 (3): 678–699.

Clerici, R., Girald, A., and S. Meggiolaro. 2015. "The determinants of academic outcomes in a competing risks approach: evidence from Italy." *Studies in Higher Education* 40 (9): 1535–1549.

Craft, J. 2019. "Academic performance of low SES students at an Australian university satellite campus." *Studies in Higher Education* 44 (8): 1372–1385.

Dahlgren, L., Fejes, A., Abrandt-Dahlgren, M., and N. Trowald. 2009. "Grading systems, features of assessment and students' approaches to learning." *Teaching in Higher Education* 14 (2): 185–194.

DeBacker, T., Crowson, H., Beesley, A., Thoma, S., and N. Hestevold. 2008. The challenge of measuring epistemic beliefs: An analysis of three self-report instruments. *Journal of Experimental Education* 76 (3): 281–312.

DeBerard, M., Spielmans, G., and D. Julka. 2004. "Predictors of academic achievement and retention among college freshman: A longitudinal study." *College Student Journal* 38 (1): 66–80.

De Clercq, M., Galand, B., and M. Frenay. 2017. "Transition from high school to university: A person-centered approach to academic achievement." *European Journal of Psychology of Education* 32 (1): 39–59

Diseth, A. 2002. "The relationship between intelligence, approaches to learning and academic achievement." *Scandinavian Journal of Educational Research* 46 (2): 219–30.

Diseth, A., Pallesen, S., Brunborg, G., and S. Larsen. 2010. "Academic achievement among first semester undergraduate psychology students: The role of course experience, effort, motives and learning strategies." *Higher Education* 59: 335–352.

Duff, A., Boyle, E., Dunleavy, K. and J. Ferguson. 2004. "The Relationship between Personalities, Approach to Learning and Academic Performance." *Personality and Individual Differences* 36: 1907–1920.

Dupont, S. Galand, B., and F. Nils. 2015. "The impact of different sources of social support on academic performance: Intervening factors and mediated pathways in the case of master's thesis." *European Review of Applied Psychology* 65 (5): 227–237.

Entwistle, N. 2009. *Teaching for understanding at university: Deep approaches and distinctive ways of thinking*. Palgrave Macmillan (Chapter 1).

Entwistle, N., and V. Mac Cune. 2013. “The disposition to understand for oneself at university: Integrating learning processes with motivation and metacognition.” *British Journal of Educational Psychology* 83 (2): 267–279.

Fokkens-Bruinsma, M., Vermue, C., Deinum, J-F., and E. Rooij. 2021. “First-year academic achievement: the role of academic self-efficacy, self-regulated learning and beyond classroom engagement.” *Assessment and Evaluation in Higher Education* 46 (7): 1115–1126.

Gijbels, D., de Watering, G. V., Dochy, F., and P. den Bossche. 2005. “The relationship between students’ approaches to learning and the assessment of learning outcomes.” *European Journal of Psychology of Education* 20 (4): 327–341.

Greco, A., Annovazzi, C., Palena, N., Camussi, E., Rossi, G., and P. Steca. 2022. “Self-efficacy beliefs of university students: Examining factor validity and measurement invariance of the new academic self-efficacy scale.” *Frontiers in Psychology 12*: Article 498824.

Guo, J-P., Yang, L-Y., Zhang, J., and Y-J. Gan. 2022. “Academic self-concept, perceptions of the learning environment, engagement, and learning outcomes of university students: Relationships and causal ordering.” *Higher education* 83, 809–828.

Hellas, A., Ihantola, P., Petersen, A., Ajanovski, V., Gutica, M., Hynninen, T., Knutas, A., Leinonen, J., Messom, C., and S. Nam Liao. 2018. “Predicting academic performance: A systematic literature review.” In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education* (pp. 175–199).

Hofer, B., and P. Pintrich. 1997. “The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning.” *Review of Educational Research* 67 (1): 88–140.

Khishfe, R. 2023. Improving students’ conceptions of nature of science: A review of the literature. *Science & Education* 32: 1887–1931.

Lehmann, T. 2022. “Student teachers’ knowledge integration across conceptual borders: the role of study approaches, learning strategies, beliefs, and motivation.” *European Journal of Psychology of Education* 37: 1189–1216.

Lin, Y.-C., Liang, J.-C., and C.-C. Tsai. 2012. “The relationships between epistemic beliefs in biology and approaches to learning biology among biology-major university students in Taiwan.” *Journal of Science Education and Technology* 21: 796–807.

Lin, T.-J., and C.-C. Tsai. 2013. “A multi-dimensional instrument for evaluating Taiwan high school students’ learning self-efficacy in relation to their approaches to learning science.” *International Journal of Science and Mathematics Education* 11 (6): 1275–1301.

Liu, E., Ye, C. and D. Yeung. 2015. “Effects of approach to learning and self-perceived overall competence on academic performance of university students.” *Learning and Individual Differences* 39: 199–204.

Lonka, K., Ketonen, E., and J.-D. Vermunt. 2021. “University students’ epistemic profiles, conceptions of learning, and academic performance.” *Higher education* 81: 775–793.

Macaulay, L., Webber, W., and Fraunholz, B. 2023. “Successful university students from low socio-economic backgrounds’ perspectives on their academic success: A capital-based approach.” *Higher Education Research & Development* 42 (8): 1986–1999.

Mattick K, Dennis I and J. Bligh. 2004. “Approaches to learning and studying in medical students: validation of a revised inventory and its relation to student characteristics and performance.” *Medical Education* 38 (5): 535–43.

Minbashian, A., Huon, G., and K. Bird. 2004. “Approaches to studying and academic performance in short-essay exams.” *Higher Education* 47: 161–176.

Mouratidis, A., Michou, A., Sayil, M., and S. Altan. 2021. “It is autonomous, not controlled motivation that counts: Linear and curvilinear relations of autonomous and controlled motivation to school grades.” *Learning and Instruction* 73: Article 101433.

Musso, M., Hernández, C., and E. Cascallar. 2020. “Predicting key educational outcomes in academic trajectories: A machine-learning approach.” *Higher Education* 80: 875–894.

Nelson Laird, T., Shoup, R., Kuh, G., and M. Schwarz. 2008. “The effects of discipline on deep approaches to student learning and college outcomes.” *Research in Higher Education* 49: 469–494.

Paivandi, S. 2015. *Apprendre à l’université*. De Boeck Supérieur (Chapter 1).

Pinxten, M., De Fraine, B., Van Den Noortgate, W., Damne, J., Boonen, T., and G. Vanlaar. 2015. “‘I choose so I am’: a logistic analysis of major selection in university and successful completion of the first year.” *Studies in higher education* 40 (10): 1919–1946.

Ramsden, P. 2003. *Learning to teach in higher education*. London & New York: Routledge.

Rania, N., Siri, A., Bagnasco, A., Aleo, G., and S. Loredano. 2014. “Academic climate, well-being and academic performance in a university degree course.” *Journal of Nursing Management* 22: 751–760.

Román, S., Cuestas, P., and P. Fenollar. 2008. “An examination of the interrelationships between self-esteem, others’ expectations, family support, learning approaches and academic achievement.” *Studies in Higher Education* 33 (2): 127–138.

Rosenbusch, H., Soldner, F., Evans, A., and M. Zeelenberg. 2021. “Supervised machine learning methods in psychology: A practical introduction with annotated R code.” *Social and Personality Psychology* 15 (2): Article e12579.

Ryan, R., and E. Deci. 2000. “Intrinsic and extrinsic motivations: Classic definitions and new directions.” *Contemporary Educational Psychology* 25 (1): 54–67.

Ryan, R., and E. Deci. 2020. “Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions.” *Contemporary Educational Psychology* 61: Article 101860.

Schiefer, J., Edelsbrunner, P., Bernholt, A., Kampa, N., and A. Nehring. 2022. “Epistemic beliefs in science: A systematic integration of evidence from multiple studies.” *Educational Psychology Review 34*: 1541–1575.

Schommer, M. 1990. “Effects of beliefs about the nature of knowledge on comprehension.” *Journal of educational psychology* 82 (3): 498–504.

Schrempft, S., Piumatti, G., Gerbase, M. and A. Baroffio. 2021. “Pathways to performance in undergraduate medical students: role of conscientiousness and the perceived educational environment.” *Advances in Health Sciences Education* 26: 1537–1554.

Sothan, S. 2019. “The determinants of academic performance: Evidence from a Cambodian University.” *Studies in higher education* 44 (11): 2096–2111.

Suphi, N. and H. Yaratan. 2012. “Effects of learning approaches, locus of control, socio-economic status and self-efficacy on academic achievement: A Turkish perspective.” *Educational Studies* 38 (4): 419-431.

Trautwein, C. and E. Bosse. 2017. “The first year in higher education: Critical requirements from the student perspective.” *Higher Education* 73: 371–387.

Vallerand, R., Blais, M., Brière, N. and L. Pelletier. 1989. “Construction et validation de l’échelle de Motivation en Education (EME).” *Canadian Journal of Behavioural Science* 21 (3): 323–349.

Van der Zanden, P., Denessen, A., Cillessen, A. and P. Meijer. 2019. “Patterns of success: First-year student success in multiple domains.” *Studies in Higher Education* 44 (11): 2081–2095.