Identifying science students at risk in the first year of higher education : the incremental value of non-cognitive variables in predicting early academic achievement

Identifying science students at risk in the first-year of higher education: the incremental value of non-cognitive variables in predicting early academic achievement

Abstract

Science students' study success rates in the first year of higher education (FYHE) are problematic. Although a considerable amount of previous research has been carried out to investigate the determinants of students' academic achievement in FYHE, there has been little discussion about the incremental value of non-cognitive factors over and above cognitive determinants in the prediction of *early* (after the first semester of FYHE) academic achievement in a science educational context. Furthermore, the complex nature of the relationships between determinants of academic achievement is frequently neglected. An investigation that addresses these gaps is important to provide the insights necessary to identify at-risk science students early in their academic career. Therefore, the main aim of this research was to examine the incremental value of non-cognitive variables (*processing strategies, regulation strategies, academic motivation, self-concept, self-efficacy*) in predicting students' early academic achievement in a science FYHE context, over and above domain-specific prior knowledge (cognitive) and after controlling for gender, age and prior education. Hereto, path-analyses were used on the data of 781 first-year students within a faculty of science of a Belgian university college in the academic year 2016-2017. Results show that cognitive variables and pre-entry characteristics predict early academic achievement. However, after controlling for these characteristics, evidence for the assumption that non-cognitive variables are determinants of early academic achievement in science education contexts, could not be found in this study. Implications for theory and research are discussed.

Key words:

Higher education, first-year students, cognitive predictors, non-cognitive predictors, early academic achievement

Introduction

Education in the STEM fields (Science, Technology, Engineering and Mathematics) is widely acknowledged as especially important for fostering innovation and economic growth in the current information and technology age (OECD 2017). Nevertheless, in most countries, these study disciplines are less popular among students and attrition rates in higher education remain a compelling problem (Chen 2013; OECD 2008; OECD 2017). This has resulted in an increased research interest in the determinants of study success in STEM (e.g. Pinxten et al. 2017; Van Soom and Donche 2014) and numerous initiatives to improve STEM students' academic achievement (e.g. Canning et al. 2017; Flemish Government 2012).

In this context, an increasing number of higher education institutions are developing interventions and practices to guide and support at-risk science students in their first year of higher education (FYHE) (Brooman and Darwent 2013; du Preez and McGhie 2015; Hultberg et al. 2008). Several lines of research establish the importance of these guidance and support initiatives early on in the academic year (e.g. Beck and Davidson 2001; Pistilli and Arnold 2010). Experiences in the early stages of higher education have been found crucial to students' adjustment and, consequently, to their long-term experiences and academic success (Harackiewicz et al. 2002; Woosley, 2003). Moreover, it has been argued that when students realise a risk of failure after receiving multiple unsatisfactory grades, the likelihood that any change in study behaviour will lead to improved performance decreases (Jayaprakash et al. 2014). Put differently, early study failure can induce a downward spiral of low self-esteem, motivation and/or academic disengagement (Wigfield et al. 2006), ultimately leading students away from the STEM field (Gasiewski et al. 2011).

It is difficult for educators of FYHE to monitor students' academic progress. Consequently, struggling students are often overlooked (Macfadyen and Dawson 2010; Yorke 2000). In this light, the development of warning systems that flag possible at-risk science students early in their academic career - before low grades are attained or adjustment problems jeopardize their progress in higher education - seems imperative for their smooth transition into higher education (Pistilli and Arnold 2010). Therefore, a predictive model containing a diversity of relevant determinants of early (after the first semester) academic achievement could be helpful in identifying at-risk students early in higher education, so that they can be led to the appropriate coaching and guidance initiatives in time.

However, the vast majority of previous research into the determinants of first-year students' academic achievement has focused on academic achievement after the completion of the first academic year (e.g. Richardson et al. 2012). In what follows, this research is explored and a number of acknowledged relationships between academic achievement and several cognitive, non-cognitive and pre-entry characteristics are described.

Determinants of academic achievement in FYHE

1. Cognitive factors

It is widely accepted that cognitive ability - in this study further addressed as prior knowledge - is an important predictor of students' academic achievement in FYHE. Echoing Dochy et al. (1999) prior knowledge is viewed as "the whole of a person's actual knowledge that is: (a) available before a certain learning task, (b) structured in schemata, (c) declarative and procedural, (d) partly explicit and partly tacit, and (e) dynamic in nature and stored in the knowledge base" (p.146). Thompson and Zamboanga (2004) found that prior knowledge of psychology shows a positive correlation with academic achievement in an introductory psychology course. Fonteyne et al. (2015) investigated a test of basic mathematics skills, taken at the beginning of FYHE, can contribute to the prediction of general academic achievement. The results of their study indicate that such a test indeed helps identifying at-risk students early in the academic year. In the STEM-context then, Ackerman et al. 2013 and Le et al. 2014 found that students' prior knowledge, conceptualised as (math) scores on matriculation exams (SAT and ACT) at the end of secondary education and high school grade point average (GPA), is positively related to academic achievement in the first year of a STEM program. Furthermore, a review study by Richardson et al. (2012) confirms a strong, positive relationship between prior knowledge and students' academic achievement across a variety of study disciplines in FYHE. Dochy et al. (2002) even posit that prior knowledge is the most important student variable in predicting learning results.

2. Non-cognitive factors

Next to the irrefutable role of prior knowledge in explaining first-year academic achievement, it has been argued that also non-cognitive factors such as meta-cognitive and motivational variables are important to include in models predicting academic achievement of students in FYHE (e.g. Credé and Kuncel 2008; Richardson et al. 2012). Indeed, Komarraju et al. (2013) state that students with low prior knowledge may rely heavily on non-cognitive psychosocial skills, such as greater solvency, dedication, and more effective study strategies to keep pace in their academic pursuits. On the other hand, of course, there are those students who achieve less well regardless of their high prior knowledge (Komarraju et al. 2013).

In the field of research on higher education students' motivation and learning, two important contrasting, yet overlapping, research traditions can be discerned (Pintrich 2004). The first tradition concerns the work by educational researchers mainly from Europe and Australia; the Student Approaches to Learning (SAL) tradition (e.g. Biggs 2003; Entwistle and Ramsden 2015; Vermunt 2005). An important theoretical framework within this SAL-perspective is Vermunt's general model of constructive learning (1998), which was substantiated in multiple studies (e.g. Vermunt and Donche 2017; Vermunt and Vermetten 2004). According to this model, processing strategies that students use are directly influenced by the regulation strategies they possess. The model further assumes that motivational factors have an

influence on processing strategies, mainly indirectly, through the use of regulation strategies. Next, the theory postulates that the way in which students regulate their learning is to a considerable degree determined by their motivation. Finally, Vermunt argues that each component in this framework has an influence on learning outcomes.

The second research tradition, the Self-regulated learning (SRL) perspective, is founded on the scientific work of predominantly North American scholars (e.g. Pintrich 2004; Zimmerman 2000). Within the SRL-perspective, Pintrich and Zusho (2007) described an important broad framework of motivation and self-regulated learning for the college classroom. This model posits that learning outcomes such as persistence and academic achievement are directly influenced by both motivational (among which self-beliefs) and self-regulatory processes. More specifically, according to Zusho (2017), the model largely assumes that students with more adaptive motivational profiles (e.g. higher self-efficacy beliefs, more highly motivated) will employ more appropriate processing and regulation strategies, and as such are more likely to attain higher levels of academic achievement.

These two authoritative models have repeatedly demonstrated their relevance for the field of higher education research on non-cognitive variables, and thus serve as the fundamental rationale for the selection of the following predictors of academic achievement: processing strategies, regulation strategies, academic self-beliefs (self-efficacy and self-concept), and academic motivation.

Research on determinants of academic achievement, however, largely concentrates on the identification of individual determinants, whereas the interplay of these factors has been less investigated (e.g. De Clercq et al. 2013a; Phan 2009; Tinto 1993, 1997). However, when exploring this previous research, the complex nature of the relationships between determinants of academic achievement becomes apparent. Therefore, an integrative research perspective considering mediating effects appears to be appropriate to capture the complexity of the interrelationships between the described influencing factors. Therefore, in accordance with the two described models, in this study we assume that (1) academic motivation and academic self-beliefs have a direct influence on processing strategies and regulation strategies; and (2) regulation strategies have a direct impact on processing strategies.

In the subsequent paragraphs, these concepts are described and the relationships within the research model (Figure 1) are further addressed.

Processing strategies

Processing strategies are those thinking activities and study skills students possess and apply to process learning contents (Vermunt 1998). In educational research, a distinction is usually made between two broad dimensions: deep and surface processing (e.g. Biggs 2003). The former is associated with thinking activities leading the student to in-depth understanding of a particular task, for instance relating or applying. The latter includes learning activities like memorising or rote learning that lead to the learning of superficial features of a study task. It should be noted that the construct of processing strategies is conceptually very similar to the construct of cognitive engagement, as detailed mainly in SRL and engagement literature (see for instance Greene et al. 1996; Greene 2015; Pintrich and Zusho 2007).

A considerable amount of previous research has shown that deep learning elicits high academic achievement (e.g. Donche et al. 2013a; Vermunt 2005; Zeegers 2001), while surface learning leads to lower academic achievement (Lindblom-Ylänne and Lonka 1998; Vermunt 2005). Moreover, the aforementioned relationships are substantiated by two meta-analyses by Richardson et al. (2012) and Dent and Koenka (2016), who found these associations to be significant, but small.

Regulation strategies

Beside processing strategies, regulation strategies play an important role in the prediction of academic achievement as well. Regulation strategies refer to those activities students use to harness their cognitive processing strategies (Schunk and Zimmerman 2012). Previous research has found self-regulation, or the extent to which students actively steer their own learning process, to be related to higher achievement (e.g. Richardson et al. 2012; Vermunt 2005). Particularly students' lack of regulation, or a lack of clarity on how to steer the learning process, has repeatedly been found associated with lower academic achievement (e.g. Busato et al. 1989; Donche and Van Petegem 2010; Van Rooij et al. 2018; Vermunt 2005).

Furthermore, previous research has confirmed the existence of interrelations among regulation and processing strategies as well. A study of Heikkilä et al. (2011), for instance, found significant relations between students' self-regulation and a deep approach to learning, while there was no relation between a self-regulated learning strategy and a surface approach to learning.

Academic motivation

Another important predictor of academic achievement distinguished in recent state of the art work is students' academic motivation. According to the self-determination theory (SDT) a distinction can be made between autonomous and controlled motivation based on the quality of the motivation (Deci and Ryan 2000). Autonomously motivated students study out of a personal interest (intrinsic regulation) or a perceived relevance for the future (identified regulation). Controlled motivation, on the other hand, refers to external sources (external regulation) or wanting to meet expectations of others as a student's

reasons to study (introjected regulation). Further, through the concept of amotivation, SDT incorporates the idea that students can lack motivation to study.

Although not all studies found significant relations between autonomous, controlled or amotivation and college students' academic achievement (e.g. Baker, 2004), in line with the numerous studies that did, we expect that the quality and quantity of motivation will be associated with academic achievement (Deci and Ryan 2000). For instance, amotivation has been repeatedly found to be negatively related with students' study success (e.g. Prospero and Vohra-Gupta 2007; Vanthournout et al. 2012). Other research shows that students who are more autonomously motivated have higher academic achievement than students who are more controlled motivated or more amotivated (e.g. Bailey and Phillips 2016; Guay et al. 2010).

Besides the relationship between academic motivation and achievement, we can also expect important relationships between first-year students' academic motivation and the processing and regulation strategies. The study of Donche et al. (2013b) demonstrated that that autonomous academic motivation was positively associated with the use of deep and - to a lesser extent - surface processing strategies. Furthermore, autonomous motivation was a positive predictor of self-regulation and a negative predictor of lack of regulation. Controlled motivation was found to be negatively associated with the use of deep processing strategies, but positively associated with the use of surface strategies. Also, in this same study, controlled motivation had a positive effect on lack of regulation, and a negative effect on self-regulation. Finally, amotivation was positively related to students' lack of regulation and – against expectations – to their self-regulation.

Self-beliefs

Over the decades, a considerable amount of studies in educational research have resorted to either self-concept or self-efficacy to conceptualize self-beliefs in educational contexts (Bong and Skaalvik 2003). Bandura (1995) describes self-efficacy as the beliefs in one's capabilities to organize and execute the courses of action required to manage prospective situations and to produce given attainments. Echoing Shavelson et al. (1976), self-concept is defined as a person's perception of himself formed through experiences with the environment. As a construct, self-concept relates strongly to the construct of self-efficacy. Zimmerman (2000), however, states that self-concept is a more general self-descriptive construct that incorporates many forms of self-knowledge and self-evaluative feelings. He argues that measures of self-concept emphasize self-esteem reactions by posing self-evaluative questions, for instance ''How good are you in English?''. Self-efficacy-items, on the other hand, focus exclusively on task-specific performance expectations, such as ''How certain are you that you can construct a correct sentence ?'' (Zimmerman 2000).

A significant body of research has found a strong positive relationship between self-efficacy and academic achievement (Richardson et al. 2012; Robbins et al. 2004). Self-concept has been found to

positively relate to academic achievement as well (e.g. Huang 2011). Despite this theoretical relation, the meta-analysis of Robbins et al. (2004) only shows a small positive relationship between the two constructs.

Self-beliefs are also important predictors of students' learning strategies. Schunk and Zimmerman (2012) conclude that self-efficacy has a strong influence on students' self-regulation. For instance, research has indeed shown that students with high self-efficacy report higher levels of lack of regulation in their learning (Donche et al. 2010). Further, studies by Fenollar et al. (2007) and Phan (2010) found a positive effect of self-efficacy on deep processing. In relation to surface processing, these studies reported inconclusive findings; Fenollar et al. (2007) found this relationship to be negative, whereas Phan (2010) reported a positive association.

3. Combination of cognitive and non-cognitive factors

Several studies have explored the value of combining cognitive and non-cognitive to explain first-year students' academic achievement. For instance, a study by Fonteyne et al. (2017) found that academic achievement after a completed academic year could be predicted by cognitive and background factors and by conscientiousness, self-efficacy and test anxiety. Even though the predictive power of variables varied across different study programs, the study concludes that the inclusion of non-cognitive factors is important and allows for a better prediction of academic achievement in several programs. Research on early warning systems in FYHE using learning analytics found that a predictive model containing solely SAT scores mildly predicted academic achievement. However, when adding students' online activity data (number of logins in a learning management system), indicating the effort exerted by a student in a course, the predictive power of the model nearly tripled (Campbell et al. 2006; Macfadyen et al. 2010). The relevance for inclusion of non-cognitive variables is also evidenced in Pinxten et al. (2017). Science students' motivation/persistence, concentration, and time management skills at the start significantly influenced student achievement at the end of the first year, although the incremental value over prior achievement was small.

Two important meta-analyses corroborate the idea that non-cognitive factors do contribute unique variance to first-year students' academic achievement next to cognitive determinants. Firstly, Credé and Kuncel (2008) uncovered that motivational variables and study skills yield substantial incremental validity in predicting academic performance in FYHE over and above both high school grades and scores on standardized admissions tests. Secondly, Richardson et al. (2012) showed that academic self-efficacy, grade goal, and effort regulation collectively accounted for 6% of the explained variance in GPA in FYHE over and above secondary school GPA and SAT/ACT scores (total model explained 28% of variance). However, to date, research on the predictive validity of the non-cognitive variables under scrutiny tended to focus on the field of humanities. Empirical evidence on this subject in the STEM-field is rather scarce.

A limited number of studies specifically explored the predictive power of a combination of cognitive and non-cognitive factors on *early* (after the first semester) academic achievement in FYHE. Harackiewicz et al. (2002) demonstrated that achievement goals, ability measures, and prior high school performance of first-semester freshmen in an introductory psychology course, each contributed unique variance in predicting academic achievement after the first semester. Olani (2009) could not find any incremental predictive validity of motivation and self-efficacy over and above cognitive variables in the total sample of 214 Ethiopian students within several study disciplines. In a STEM-context, Van Soom and Donche (2014) found that, for male students and after controlling for prior high school result, 4% of variance in early academic achievement could be explained by student profiles that were based upon autonomous motivation and academic self-concept. For female students, there was no significant effect of profile assignment on early academic achievement after controlling for high school results.

4. Pre-entry characteristics

FYHE research has demonstrated that students' pre-entry characteristics *age, gender* and *prior secondary education* are related to academic achievement (See Bruinsma and Jansen 2007; De Clercq et al. 2013b; Donche and Van Petegem 2010; Fonteyne et al. 2015; Jansen 2004, Richardson et al. 2012), and to processing and regulation strategies (see Donche and Van Petegem 2010, Severiens and Ten Dam 1997; Vermunt 2005). In this study we included these pre-entry characteristics as control variables.

This study

Although considerable research has been devoted to the identification of determinants of academic achievement in FYHE, less attention has been paid to the incremental value of a wide variety of non-cognitive factors in combination with cognitive determinants in the prediction of *early* academic achievement in science education. Furthermore, this strand of research typically neglects the complex nature of the relationships between determinants, focussing solely on the identification of individual determinants. Indeed, to our knowledge, the process model that we set forth (figure 1), containing this particular composition of abovementioned non-cognitive variables, has not yet been tested in a science FYHE context.

Therefore, this study aims to examine the incremental value of processing strategies, regulation strategies, academic motivation, self-efficacy, self-concept (non-cognitive) in predicting students' early academic achievement in a science FYHE context, over and above domain-specific prior knowledge (cognitive) and after controlling for gender, age and prior education. To this end, an integrative research perspective, considering mediating effects is adopted to capture the complexity of the interrelationships between the described influencing factors. This way, we attempt to contribute to the insights necessary to identify at-risk science students early in their academic career.

Summarizing, and based on previous research outlined in the theoretical framework, the following conceptual model was drafted (Figure 1). It depicts the hypothesised relationships between the discussed constructs that will be explored in this study. The expected directions of these relationships are further detailed in Table 1.

[Insert Figure 1 about here]

[Insert Table 1 about here]

Methodology

Sample & procedure

This study took place within two departments of one faculty of sciences of a Belgian university in academic year 2016-2017. Data of 781 first-year students, 519 of which (66,5%; 15 missing) female, were used. Seven science study disciplines were involved: *pharmaceutical sciences, biomedical sciences, veterinary sciences, biology, chemistry, bioscience engineering*, and *biochemistry and biotechnology*. The first semester of the first year of these study disciplines comprise mainly general science introductory courses. For instance, in all of the seven study programs a chemistry course is taught, which counts for five to seven credits (ECTS study points), depending on the program considered. In that vein, the seven study disciplines have a similar character. Nevertheless, the disciplines also contain several program-specific courses. For example, the course *Evolution theory and biological classification* is only taught in the biology program.

We analysed data that had previously been collected for guidance purposes. Two measurement moments were used. After each measurement moment, participating students received an individual e-mail with quantitative and qualitative (textual) feedback that was based upon their test scores. Furthermore, students were invited to talk about their results with student counsellors. During the first measurement moment (Wave1; N=731) at the start of the first semester, online surveys regarding the motivational learning components and tests measuring prior knowledge were made accessible for all students. Processing and regulation strategies were mapped at the second measurement moment (Wave2; N=592), in the first week of December (two months after the start of the first semester). All students in the seven study disciplines were informed that participation was mandatory as a part of their study trajectory. However, non-completion of the tests and questionnaires did not result in any form of penalty.

Measurements

With the exception of the construct *self-concept*, all *non-cognitive* variables in this study were measured using scales of the validated self-report questionnaire 'LEarning strategy and MOtivation questionnaire (LEMO; Donche and Van Petegem 2008; Donche et al. 2010). Self-efficacy in LEMO is defined more specifically as a student's perception of having the necessary knowledge and skills to carry out a particular learning task. Further, a three items scale was used to measure the construct self-concept, which was conceptualised as academic self-beliefs regarding the chosen discipline. The reliability and validity of this scale for students in FYHE has previously been established (Van Soom and Donche 2014). Next to the aforementioned concepts, three *motivational* scales (autonomous motivation, controlled motivation, amotivation), two *processing strategies* scales (deep processing, surface processing) and two *regulation strategies* scales (self-regulation, lack of regulation) were included in this study as well. Table 2 shows the number of items, the measurement moment, an item example and the reliability of all non-cognitive constructs under study.

[Insert Table 2 about here]

Prior knowledge (Wave1). Several years ago, the two departments of science of the university wherein this study took place, introduced orientation tests at the beginning of the academic year. These tests explore the extent to which students have the necessary knowledge of mathematics and chemistry to start the different scientific study programs. In the present study, students' scores on these tests will be used to conceptualise their domain-specific prior knowledge of mathematics and chemistry. The inclusion of chemistry as the prior knowledge domain, instead of other sciences, was mainly based on the argument that chemistry is important subject matter in the first semester of all of the seven study programs (for instance, physics is not). Moreover, the more pragmatic argument that there already existed a chemistry orientation test further justified this conceptualisation.

The mathematics test consists of 20 multiple-choice questions (α=.76), of which "In which points of the graph of the function $f(x) = -x^2 + x - 6$ does the tangent have slope -1?" is an example. The chemistry test comprises 30 multiple-choice questions (e.g. "Which one of the following solutions is a buffer solution?", α=.80). These tests have been drafted by specialists in the respective fields of study and have been optimized in the last couple of years, based upon several reliability analyses. Exploratory factor analysis showed that, for both tests, all items loaded on one factor.

Pre-entry characteristics. Data on students' *prior education*, *gender* and *age* were gathered through the use of a self-report questionnaire. For further use in the analysis process, the *gender* variable was dummy coded (0=female, 1=male). The indicator for *prior education* in this study are the number of hours of

mathematics, chemistry and physics students received weekly in their last year of secondary education. These variables were again dummy coded: maths (0=four hours or less, 1=five hours or more), physics (0=one hour or less, 1=two hours or more), and chemistry (0=one hour or less, 1=two hours or more). Students' *age* was conceptualised by making a division into two categories: students who are 18 or younger (code=0) and students who are older than 18 (code=1; 29.1% of respondents).

Early academic achievement. Data concerning academic achievement of respondents were obtained through administration databases of the faculty involved, after the first examination period in January. Early academic achievement was conceptualised as study progress; the number of credits (ECTS study points) obtained by a student at the end of the first semester, relatively to the total number of credits registered for in that semester by the concerning student (ratio credits earned/credits attempted). A student passed a course (earns credits), when he or she scored a minimum of 10 out of 20 on the evaluation for that course.

Analysis

To answer the research question in this study, structural equation modelling (SEM) was applied, using softwarepackage *lavaan* in R (Rosseel 2012). This technique allows the modelling of the indirect relationships between the constructs under study. Moreover, structural equation models allow to model the variables as latent variables, thus filtering out measurement error (Kline 2016).

The complex structure of the data (respondents are 'nested' within seven study disciplines) would suggest a multilevel approach (Gelman and Hill 2006). Indeed, when the clustering of observations is ignored, this can lead to an increased likelihood of committing Type I errors (Thomas and Heck 2001). Maas and Hox (2005), however, state that sample size at level two needs to be sufficiently large to be able to accurately estimate the second-level standard errors. Even merely controlling for complex data structures, would require at least 20 clusters (Muthén 2005). Due to the fact that in the present study the number of units at level two (i.e., seven) is less than 20, the model was fit with SEM at individual level. Additionally, further exploration of associations between subjects within the clusters, using multiple-group SEM analysis, deemed to be impossible since the sample sizes within the single study disciplines (e.g. smallest: biology, N=46) was too limited.

As the data contain missing values, which is a common problem in studies using multiple measurement moments (Engels and Diehr 2003); a maximum likelihood missing data handling procedure was used, also called full information maximum likelihood (FIML), which gives unbiased results when missing values are missing at random, as was assumed (Wang and Wang 2012). FIML is considered to be superior to traditional techniques (e.g., listwise deletion) because it maximizes statistical power by borrowing information from the observed data. This means that maximum likelihood will produce accurate parameter estimates in situations where more traditional approaches cannot (Enders 2010).

In general, two steps can be discerned in the estimation of SEM-models (Schumacker and Lomax 2010). In a first step of the analysis, the measurement model is specified to define all relationships between the latent and the manifest variables. In this step, also called 'confirmatory factor analysis', the factor structure of the different latent variables was verified.

In a second step of the analysis, the structural model was specified to investigate the relationships between all latent variables. To compare the goodness of fit of different models, first of all the Akaike Information Criterion (AIC) was used. A lower value on this index reflects better fit of one model to the data as compared to a model with a higher AIC-value (Kline 2016). Next to AIC, three other indices were used to identify the fit of different models: 'comparative fit index' (CFI), 'root mean square error of approximation' (RMSEA), and 'standard root mean square' (SRMR). Hu and Bentler (1999) state that a model has excellent fit when CFI has a value above .95, RMSEA has a value less than .05 and SRMR has a value less than .06; a model has acceptable fit, with a CFI-value above .90, and RMSEA and SRMR values less than .08.

Results

As a first step in the modelling of the data, confirmatory factor analysis was performed to investigate to what extent the measurement model has good construct validity. After adding three error covariance terms, a model containing all nine 'non-cognitive components' was obtained that fits the data well. The fit indices indicated that this model (see Appendix A for more details) has an acceptable (*CFI*=.92) to excellent fit (*RMSEA*=.04 and *SRMR*=.06).

Subsequently, to get a first grasp on the relationships between the variables in the structural part of the SEM-models, Pearson product Moment Correlations among all constructs under study were calculated (Table 3). The table shows moderate to strong (Cohen 1988) significant correlations between early academic achievement on the one hand, and the number of hours of mathematics in the last year of secondary education (*r*=.25), age (*r*=-.23), domain-specific prior knowledge of mathematics (*r*=.46) and chemistry (*r*=.46) on the other hand. Further, early academic achievement is weakly correlated with self-efficacy (*r*=.17), self-concept (*r*=.13), deep processing (*r*=.09), lack of regulation (*r*=-.08), amotivation (*r*=-0.09), and the number of hours of physics in the last year of secondary education (*r*=.17). Finally, several significant weak to strong correlations are uncovered between the motivational and metacognitive factors and pre-entry characteristics.

In a next step, SEM was applied to further investigate the relationships among these factors and early academic achievement. Based on previous research outlined in the theoretical framework - and depicted in Figure 1 - an initial baseline model, containing all theoretical paths, was tested. Although the fit indices of this model (*CFI*=.90, *RMSEA*=.04, *SRMR*=.06, *AIC*=70320.765) were acceptable to excellent, several paths were not significant and it was predicted that a revised model could result in a better fit.

A second phase in the specification of our model's structural part, comprised the estimation of a modified and more parsimonious model by considering only significant paths (Byrne, 2010). Furthermore, in this process, we tested several alternative models, considering different patterns of direct effects between non-cognitive variables and the academic achievement outcome (See appendix B). The fit indices of the resulting model predicting early academic achievement (Figure 2) were: *CFI*=.91, *RMSEA*=.04, *SRMR*=.06, *AIC*=70310.011). Based upon these fit indices, we concluded that the resulting path model has better fit with the data than the baseline model. In what follows, the results of this SEM-analysis are discussed.

[Insert Table 3 about here]

[Insert Figure 2 about here]

First of all, a number of variables measured at wave 1 predict regulation strategies at wave 2. The results show that students that are older than 18 years old are more self-regulated in their learning (β=.167, p<.01), than students that are 18 years old or younger. Similarly, students who are more autonomously motivated tend to show higher levels of self-regulation (β=.441, p<.001). Further, it appears that students who score higher on amotivation (β=.129, p<.05) score higher on lack of regulation, while students with high self-concept (β=-.161, p<.01), students with high self-efficacy (β=-.482, p<.001) and male students (β=-.134, p<.01) have lower levels of lack of regulation. Finally, it seems that the level of controlled motivation is positively associated with students' lack of regulation (β=.183, p<.05). The analyses indicate that the variance in lack of regulation could be fairly well explained within this model (54.2%). This is in contrast to the explained variance of self-regulation, which amounts to 22.3%.

Next, the predictors of processing strategies were investigated as well. Evidence was found that students who are more able to self-regulate their own learning, make more use of the processing strategy 'Relating & Structuring' which is a deep processing strategy (β=.565, p<.001). Students who report a high level of lack of regulation on the other hand, less often use a deep processing strategy (β=-.367, p<.001). The surface processing strategy 'Memorizing' was found to be related with self-regulation, self-efficacy and controlled motivation. Students that report higher self-efficacy beliefs at the first wave, have higher scores on the surface processing strategy at the second wave (β=.197, p<.01). The results further show that students who are more controlled motivated have a more superficial learning style at the second wave (β=.177, p<.01). Finally, it appears that students who rely more on self-regulatory strategies will use more superficial learning strategies (β=.195, p<.01). The explained variance of deep processing in this model is 51.9%, while only 9.2% of the variance of surface processing was predicted.

Predictors of early academic achievement

The path model presented in Figure 2 shows that early academic achievement of students is predicted by several pre-entry factors. For instance, students who are 19 years old or older obtain a smaller percentage of their adopted credits than students that are 18 years old or younger (β=-.171, p<.001). Students that had received weekly more than four hours of mathematics in the last year of their secondary education, have higher early academic achievement than students receiving four hours or less (β=.129, p<.001). Further, early academic achievement of men is lower than that of woman (β=-.118, p<.001). The results also indicate that domain-specific prior knowledge of mathematics (β=.162, p<.01) and chemistry (β=.284, p<.001) predicts students' early academic achievement. It is noteworthy that most hypothesised direct relations between meta-cognitive and motivational predictors and students' early academic achievement are absent. Only lack of regulation seems to have a negative direct effect on this achievement measure (β=-.123, p<.01). As stated above, lack of regulation is in turn predicted by several motivational constructs, entailing various indirect effects on academic achievement.

Table 4 shows that this complete model - containing pre-entry, cognitive and non-cognitive predictors - explains 28% of the variance in students' early academic achievement. Surprisingly, the results further indicate that only a small proportion of the variance of early academic achievement can be predicted by a model containing solely the non-cognitive factors (2.8%). Furthermore, after controlling for pre-entry and cognitive factors, no considerable amount of variance was predicted by the non-cognitive factors (0.3%), as can also be seen in table 3.

[Insert Table 4 about here]

Conclusion & Discussion

Science students' study success rates in FYHE are problematic. Despite the large body of research on the identification of determinants of academic achievement in FYHE (e.g. Richardson et al. 2012; Van Rooij et al. 2018), the incremental value of a wide variety of non-cognitive factors in combination with cognitive determinants in the prediction of early academic achievement in science education remains veiled. Moreover, this strand of research generally neglects the complex nature of the relationships between determinants, focussing solely on the identification of individual determinants. The present study innovates in adopting an integrative research perspective: we aimed to explore the incremental value of processing strategies, regulation strategies, academic motivation, self-efficacy, self-concept (within one integrated model) in predicting students' early academic achievement in a science FYHE context, over and above domain-specific prior knowledge of mathematics and chemistry.

Results indicate that early academic achievement of first year higher education students in an exact scientific educational context is predicted by domain-specific prior knowledge of mathematics and chemistry. On average, students with a greater prior knowledge in these fields of interest will obtain a higher percentage of their credits in the first semester. These findings confirm previous research that also pointed to a positive relationship between prior knowledge and academic achievement (Dochy et al. 2002; Fonteyne et al. 2015; Richardson et al. 2012). For example, Fonteyne et al. (2015) similarly found that a test that maps out first-year Psychology and Educational Sciences students' mathematical skills at the beginning of the first semester predicted general academic achievement well.

In contrast to previous research (e.g. Prospero and Vohra-Gupta 2007; Richardson et al. 2012), the path model shows only one direct relationship between non-cognitive variables and early academic achievement. Lack of regulation negatively predicts early academic achievement, which is in line with former research (Busato et al. 1989; Donche and Van Petegem 2010; Van Rooij et al. 2018; Vermunt 2005). Students who have no clear understanding of how to steer their own learning process will obtain a lower percentage of their credits in January.

A strength of this study is that this relationship was found after controlling for prior knowledge and pre-entry characteristics. In addition, this study went beyond the consideration of separate direct effects of motivational variables and learning strategies on early academic achievement, by examining their interplay in one integrated model. Indeed, in our model we could identify that lack of regulation is predicted by several motivational factors, causing some indirect effects of these variables on academic achievement. First of all, results showed that students who are more confident in their way of studying (self-efficacy) and students with a high general academic self-esteem (self-concept) report lower levels of lack of regulation, which is consistent with findings of previous research (Donche et al. 2010b). Secondly, as reported by previous research (Donche et al. 2013b), the resulting model indicated that students for whom the incentive to study is determined by external sources (controlled motivation), and students who score higher on amotivation at the start of the first year, scored higher on the lack of regulation scale after two months in HE.

When interpreting these results, it is important to keep in mind that the calculation of effect sizes showed that - after controlling for pre-entry and cognitive factors that together predict 27.7% of variance in early academic achievement - no substantial portion of variance in the outcome measure was predicted by the non-cognitive variables (0.3%). This finding, however, is contrary to that of several studies that found non-cognitive factors to make a significant contribution in the prediction of students' academic achievement, even after controlling for cognitive and pre-entry factors (Campbell et al. 2006; Fonteyne et al. 2017; Credé and Kuncel 2008; Pinxten et al. 2017; Richardson et al. 2012).

It should be noted that the aforementioned studies considered academic achievement after a completed academic year, while the present study examined *early* academic achievement. Previous research

focussing on this early academic achievement outcome measure report mixed findings on this matter. Our results accord with the findings of Olani (2009), who did not detect any incremental predictive value of non-cognitive variables. Harackiewicz et al. (2002) and Van Soom and Donche (2014), however, found that non-cognitive variables were able to contribute unique variance in predicting academic achievement after the first semester over and above cognitive variables.

In addition to the aforementioned, the results of the present study show that the predictive model solely containing non-cognitive factors only explains a small portion of variance in the outcome measure (2.8%), and correlations between non-cognitive variables on the one hand and early academic achievement on the other hand were weak at best. These results are contradictory to the findings of previous studies (including also other disciplines than science disciplines), that found stronger correlations (e.g. Richardson et al. 2012) and were generally able to explain a more substantial portion of the variance of academic achievement (e.g. Prospero and Vohra-Gupta 2007). Again, these studies included academic achievement at the end of a completed academic year as the achievement-outcome-measure, which further corroborates the idea that non-cognitive factors may be less important in predicting early academic achievement in science FYHE contexts.

In conclusion, cognitive factors (prior knowledge) are important to take into account when predicting early academic achievement of students in science study disciplines. However, we were unable to confirm the hypothesis that also non-cognitive variables have a considerable incremental value in predicting early academic achievement of these students. Finally, according to our results, future research should pay attention to pre-entry factors such as age, gender and number of hours of mathematics in secondary education (prior education) when predicting students' early academic achievement in FYHE.

Hypotheses, perspectives and limitations

We present here two hypotheses that may explain the limited influence of non-cognitive variables on early academic achievement. A first reason could be found in the fact that, in this study, we used learning outcomes of the examinations in January as measures of early academic achievement. However, the science study disciplines incorporated in this study comprise mainly general science introductory courses in the first semester of the first year; and we could, therefore, contemplate that the learning contents in this first semester are to a great extent recurrences of the curriculum of the last year of secondary education. In this light, it is possible that in this phase of their education it matters less which learning strategies students use, or how they are motivated, because a great deal of the required knowledge is already acquired in their prior education.

A second hypothesis concerns the extent to which the exams after the first semester sufficiently capitalise on students' use of deep and self-regulated processing strategies. Conceivably these evaluations might predominantly require students' factual knowledge (i.e. definitions and formulas) to

succeed; students' insight and deeper or critical understanding of subject matter may become more important during the second semester. This reasoning potentially brings to light a mechanism that also partially explains the limited influence of non-cognitive variables on early academic achievement, since a more adapted learning strategy is then not rewarded by the form of evaluation. Future study designs that (1) take into account both early academic achievement and academic achievement after the second semester and (2) ensure that the used methods of evaluation account for more adapted learning strategies, could create more clarity in these matters.

A limitation of the present study is that individual student factors were considered as determinants of early academic achievement, while factors related to the learning environment were not taken into account. However, previous research has shown that factors related to the learning environment are important predictors of students' academic achievement as well (e.g. Bruinsma and Jansen 2007; Ruiz-Gallardo et al. 2011). Future research should therefore also address the predictive power of determinants of early academic achievement at the institutional level, such as classroom environment, quality and quantity of instruction, teaching method and student workload.

Further, it should be addressed that academic achievement was operationalized by the ratio of credits earned/credits attempted. Although we believe that this outcome measure is an important indicator of the early study progress a student makes in FYHE, it says little about the *quality* of success. Studies incorporating GPA as achievement measure, for instance, would be more favourable if this quality was the prime research focus.

Finally, this study was limited to the use of self-report questionnaires and cognitive tests as methods of data collection. However, research in computer mediated learning suggests that more direct measures of non-cognitive determinants (e.g. online activity data indicating the effort exerted by a student in a course) may have important incremental value in predicting students' (early) academic achievement (Campbell et al. 2006; Macfadyen et al. 2010). This would be a fruitful area for further research.

Implications for practice

Notwithstanding the limitations as outlined above, our conclusions entail different practical implications. First of all, a call for educational designers to be *attentive* to science students with low levels of prior knowledge seems appropriate. In this light, the present study is only one step in the development of a model that takes into account a multitude of relevant factors in the prediction of early academic progress. Such a model could, for instance, be a great aid in *identifying* at-risk science students early in their transition into higher education. Based upon the results of this study, we state that prior knowledge should play an essential role in such an instrument (e.g. discipline specific entry-assessment). Further, it could be recommended to extend the *feedback initiatives* given to students at the beginning of higher education, based on the information about students' prior knowledge. This could lead them to the necessary *coaching* and *guidance* initiatives many institutions organise in the first months of higher

education, aimed at facilitating the transition in first year higher education. A widely known and well-researched academic support intervention in this context, is Supplemental Instruction (SI) or Peer-Assisted Study Sessions (PASS) (Dawson et al. 2014; van der Meer et al. 2017), where upper year students support first-year students in their transition process (e.g. facilitating discussions around course content, preparing learning activities such as worksheets, group work and problem-solving exercises). This is just one example of a support initiative that might help struggling science students in FYHE with prior knowledge deficits to overthrow their cognitive detriment in a specific study domain.

References

Ackerman, P. L., Kanfer, R., & Beier, M. E. (2013). Trait complex, cognitive ability, and domain knowledge predictors of baccalaureate success, STEM persistence, and gender differences. *Journal of Educational Psychology*, *105*(3), 911-927.

Bailey, T. H., & Phillips, L. J. (2016). The influence of motivation and adaptation on students' subjective well-being, meaning in life and academic performance. *Higher Education Research & Development*, *35*(2), 201-216.

Baker, S. R. (2004). Intrinsic, extrinsic, and amotivational orientations: Their role in university adjustment, stress, well-being, and subsequent academic performance. *Current Psychology*, *23*(3), 189-202.

Bandura, A. (1995). *Self-efficacy in changing societies*. New York: Cambridge university press.

Beck, H. P., & Davidson, W. D. (2001). Establishing an early warning system: Predicting low grades in college students from survey of academic orientations scores. *Research in Higher Education*, *42*(6), 709-723.

Biggs, J. (2003). *Teaching for Quality Learning at University* (2nd Edition). Buckingham: SRHE and Open University Press.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really?. *Educational psychology review*, *15*(1), 1-40.

Bruinsma, M., & Jansen, E. P. W. A. (2007). Educational productivity in higher education: An examination of part of the Walberg educational productivity model. *School Effectiveness and School Improvement*, *18*(1), 45-65.

Brooman, S., & Darwent, S. (2013). Measuring the beginning: a quantitative study of the transition to higher education. Stud High Educ, 39(9), 1523–1541. doi:10.1080/03075079.2013.801428.

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (1998). Learning styles: a cross-sectional and longitudinal study in higher education. *British Journal of Educational Psychology*, *68*(3), 427-441.

Byrne, B. M. (2010). *Structural Equation Modeling with AMOS, basic concepts, applications, and programming (2nd edition)*, New York: Routledge Taylor & Francis Group.

Canning, E. A., Harackiewicz, J. M., Priniski, S. J., Hecht, C. A., Tibbetts, Y., & Hyde, J. S. (2017). Improving Performance and Retention in Introductory Biology With a Utility-Value Intervention.

Journal of Educational Psychology. Advance online publication. http://dx.doi.org/10.1037/edu0000244

Campbell, J. P., Finnegan, C., & Collins, B. (2006, July). *Academic analytics: Using the CMS as an early warning system*. Paper presented at the WebCT impact conference 2006.

Chen, X. (2013). *STEM attrition: College students' paths into and out of STEM fields (NCES 2014-001).* Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Cohen, J. (1988). *Statistical Power Analysis for the behavioural sciences* (2nd edition). New Jersey: Lawrence Erlbaum Associates.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on psychological science*, *3*(6), 425-453.

De Clercq, M., Galand, B., Dupont, S., & Frenay, M. (2013b). Achievement among first-year university students: an integrated and contextualised approach. *European journal of psychology of education*, *28*(3), 641-662.

De Clercq, M., Galand, B., & Frenay, M. (2013a). Chicken or the egg: Longitudinal analysis of the causal dilemma between goal orientation, self-regulation and cognitive processing strategies in higher education. *Studies in educational evaluation*, *39*(1), 4-13.

Dawson, P., van der Meer, J., Skalicky, J., & Cowley, K. (2014). On the effectiveness of supplemental instruction: a systematic review of supplemental instruction and peer-assisted study sessions literature between 2001 and 2010. Review of Educational Research, 84(4), 609–639.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, *11*(4), 227-268.

Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: a meta-analysis. *Educational Psychology Review*, *28*(3), 425-474.

Dochy, F., De Rijdt, C., & Dyck, W. (2002). Cognitive prerequisites and learning. How far have we progressed since Bloom? Implications for educational practice and teaching. *Active Learning in Higher Education*, *3*(3), 265–284.

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of educational research*, *69*(2), 145-186.

Donche, V., Coertjens, L., van Daal, T., De Maeyer, S., & Van Petegem, P. (2013a). Understanding differences in student learning and academic achievement in first year higher education. An integrated research perspective. In D. Gijbels, V. Donche, J.T.E. Richardson & J.D. Vermunt

(Eds.), *Learning Patterns in Higher Education: Dimensions and Research Perspectives* (pp. 214-231). Londen/New York: Routledge.

Donche, V., De Maeyer, S., Coertjens, L., Van Daal, T., & Van Petegem, P. (2013b). Differential use of learning strategies in first-year higher education: The impact of personality, academic motivation, and teaching strategies. *British Journal of Educational Psychology*, *83*(2), 238-251.

Donche, V., & Van Petegem, P. (2008). The validity and reliability of the short inventory of learning patterns. In E. Cools, H. Van den Broeck, C. Evans, & T. Redmond (Eds.), *Style and cultural differences: How can organisations, regions and countries take advantage of style differences* (pp. 49-59). Gent, Belgium: Vlerick Leuven Gent Management School.

Donche, V., & Van Petegem, P. (2010). The relationship between entry characteristics, learning style and academic achievement of college freshmen. In M. Poulson (Ed.), *Higher education: Teaching, internationalisation and student issues*, (pp. 277-288). New York: Nova Science Publishers.

Donche, V., Van Peteghem, P., & Van de Mosselaer, H. (Eds.). (2010*). LEMO: een instrument voor feedback over leren en motivatie* [LEMO: an inventory for feedback on learning and motivation]. Mechelen: Plantyn.

du Preez, M., & McGhie, V. M. V. (2015). Addressing the learning needs of at-risk students at a South African university. *South African Journal of Higher Education*, *29*(1).

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Press.

Engels, J. M., & Diehr, P. (2003). Imputation of missing longitudinal data: a comparison of methods. *Journal of clinical epidemiology*, *56*(10), 968-976.

Entwistle, N., & Ramsden, P. (2015). *Understanding Student Learning (Routledge Revivals)*. New York: Routledge.

Fenollar, P., Roman, S., & Cuestas, P.J. (2007). University students' academic performance: An integrative conceptual framework and empirical analysis. *British Journal of Educational Psychology*, *77*, 873–891.

Flemish Government (2012). *Actieplan voor het stimuleren van loopbanen in wiskunde, exacte wetenschappen en techniek* [Action plan for the stimulation of study careers in mathematics, exact sciences and techniques]. Retrieved 25 April, 2018, from https://onderwijs.vlaanderen.be /sites/default/files/atoms/files/STEM-actieplan.pdf

Fonteyne, L., De Fruyt, F., Dewulf, N., Duyck, W., Erauw, K., Goeminne, K., et al. (2015). Basic mathematics test predicts statistics achievement and overall first year academic success. *European Journal of Psychology of Education*, *30*(1), 95-118.

Fonteyne, L., Duyck, W., & De Fruyt, F. (2017). Program-specific prediction of academic achievement on the basis of cognitive and non-cognitive factors. *Learning and Individual Differences*, *56*, 34-48.

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in higher education*, *53*(2), 229-261.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.

Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist*, *50*(1), 14-30.

Greene, B. A., & Miller, R. B. (1996). Influences on achievement: Goals, perceived ability, and cognitive engagement. *Contemporary Educational Psychology*, *21*(2), 181-192.

Grewal, R., Cote, J. A., & Baumgartner, H. (2004). Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Marketing Science*, *23*(4), 519-529.

Guay, F., Ratelle, C., Roy, A., & Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: Mediating and additive effects. *Learning and Individual Differences*, *20*(6), 644-653.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, *94*(3), 562-575.

Heikkilä, A., Niemivirta, M., Nieminen, J., & Lonka, K. (2011). Interrelations among university students' approaches to learning, regulation of learning, and cognitive and attributional strategies: a person oriented approach. Higher Education, *61*(5), 513–529.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, *49*(5), 505-528.

Hultberg, J., Plos, K., Hendry, G. D., & Kjellgren, K. I. (2008). Scaffolding students' transition to higher education: parallel introductory courses for students and teachers. *Furth High Educ*, *32*(1), 47–57. doi:10.1080/03098770701781440.

Jansen, E. P. W. A. (2004). The influence of the curriculum organization on study progress in higher education. *Higher Education*, *47*(4), 411-435.

Jansen, E. P. W. A. (2004). The influence of the curriculum organization on study progress in higher education. *Higher Education*, *47*(4), 411-435.

Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, *1*(1), 6-47.

Kline, R. B. (2016). *Principles and practice of structural equation modelling*. New York: Guilford Press.

Komarraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. *Learning and Individual Differences*, *24*, 103-109.

Le, H., Robbins, S. B., & Westrick, P. (2014). Predicting student enrollment and persistence in college STEM fields using an expanded PE fit framework: A large-scale multilevel study. *Journal of Applied Psychology*, *99*(5), 915-947.

Lindblom-Ylänne, S., & Lonka, K. (1998). Individual ways of interacting with the learning environment - are they related to study success? *Learning and instruction*, *9*(1), 1-18.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *1*(3), 86-92.

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

Muthén, B. O. (2005). Example for type=complex. Retrieved 23 April 2018, from http://www.statmodel.com/discussion/messages/12/776.html?1523500311.

Olani, A. (2009). Predicting first year university students' academic success. Electronic Journal of Research in Educational Psychology, 7(3), 1053–1072.

OECD (2008), *Education at a Glance 2008: OECD Indicators*, Paris: OECD Publishing, http://dx.doi.org/10.1787/eag-2017-en.

OECD (2017), *Education at a Glance 2017: OECD Indicators*, Paris: OECD Publishing, http://dx.doi.org/10.1787/eag-2017-en

Phan, H. P. (2009). Relations between goals, self-efficacy, critical thinking and deep processing strategies: a path analysis. *Educational Psychology*, *29*(7), 777-799.

Phan, H. P. (2010). Students' academic performance and various cognitive processes of learning: an integrative framework and empirical analysis. *Educational Psychology*, *30*(3), 297-322.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational psychology review*, *16*(4), 385-407.

Pintrich, P. R., & Zusho, A. (2007). Motivation and self-regulated learning in the college classroom. In R. Perry & J. Smart (Eds.), *Handbook on Teaching and Learning in Higher Education*. Dordrecht: Springer Publishers.

Pinxten, M., Van Soom, C., Peeters, C., De Laet, T., & Langie, G. (2017). At-risk at the gate: prediction of study success of first-year science and engineering students in an open-admission university in Flanders—any incremental validity of study strategies? *European Journal of Psychology of Education*, 1-22.

Pistilli, M. D., & Arnold, K. E. (2010). In practice: Purdue Signals: Mining real-time academic data to enhance student success. *About Campus*, *15*(3), 22-24.

Prospero, M., & Vohra-Gupta, S. (2007). First generation college students: Motivation, integration, and academic achievement. *Community College Journal of Research and Practice*, 31(12), 963-975.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological bulletin*, *138*(2), 353-387.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological bulletin*, *130*(2), 261-288.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36.

Ruiz-Gallardo, J. R., Castaño, S., Gómez-Alday, J. J., & Valdés, A. (2011). Assessing student workload in Problem Based Learning: Relationships among teaching method, student workload and achievement. A case study in Natural Sciences. *Teaching and Teacher Education*, *27*(3), 619-627.

Schumacker, R., & Lomax, R. G. (2010). *A beginners guide to structural equation modeling*. New York: Routlege.

Schunk, D. H., & Zimmerman, B. J. (Eds.). (2012). *Motivation and self-regulated learning: Theory, research, and applications*. Routledge.

Severiens, S., & Ten Dam, G. T. (1997). Gender and gender identity differences in learning styles. *Educational Psychology*, *17*(1-2), 79-93.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of educational research*, *46*(3), 407-441.

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition (Second edition).* Chicago: University of Chicago Press.

Tinto, V. (1997). Classrooms as communities: Exploring the educational character of student persistence. *The Journal of Higher Education*, *68*, 599-623.

Thomas, S. L., & Heck, R. H. (2001). Analysis of large-scale secondary data in higher education research: Potential perils associated with complex sampling designs. *Research in higher education*, *42*(5), 517-540.

Thompson, R. A., & Zamboanga, B. L. (2004). Academic Aptitude and Prior Knowledge as Predictors of Student Achievement in Introduction to Psychology. *Journal of educational psychology*, *96*(4), 778-784.

van der Meer, J., Wass, R., Scott, S., & Kokaua, J. (2017). Entry characteristics and participation in a peer learning program as predictors of first-year students' achievement, retention, and degree completion. *AERA Open*, *3*(3), 1-13.

van Rooij, E., Brouwer, J., Fokkens-Bruinsma, M., Jansen, E., Donche, V., & Noyens, D. (2018). A systematic review of factors related to first-year students success in Dutch and Flemish higher education. *Pedagogische studiën*, *94*(5), 360-405.

Van Soom, C., & Donche, V. (2014). Profiling first-year students in STEM programs based on autonomous motivation and academic self-concept and relationship with academic achievement. *PloS one*, *9*(11), e112489.

Vanthournout, G., Gijbels, D., Coertjens, L., Donche, V., & Van Petegem, P. (2012). Students' persistence and academic success in a first-year professional bachelor program: The influence of students' learning strategies and academic motivation. *Education Research International*, *2012*, 1-10.

Vermunt, J. D. (1998). The regulation of constructive learning processes. *British journal of educational psychology*, *68*(2), 149-171.

Vermunt, J. D. (2005). Relations between student learning patterns and personal and contextual factors and academic performance. *Higher education*, *49*(3), 205-234.

Vermunt, J. D., & Donche, V. (2017). A learning patterns perspective on student learning in higher education: state of the art and moving forward. *Educational Psychology Review*, *29*(2), 269–299.

Vermunt, J. D., & Vermetten, Y. J. (2004). Patterns in student learning: Relationships between learning strategies, conceptions of learning, and learning orientations. *Educational psychology review*, *16*(4), 359-384.

Wang, J., & Wang, X. (2012). *Structural equation modeling: Applications using Mplus*. Chichester, United Kingdom: John Wiley & Sons.

Wigfield, A., Byrnes, J. P., & Eccles, J. S. (2006). Development during early and middle adolescence. In P. A. Alexander, & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 61–113) (2nd ed.). New York: Routledge.

Woosley, S. A. (2003). How important are the first few weeks of college? The long term effects of initial college experiences. *College Student Journal*, *37*(2), 201-208.

Yorke, M. (2000). Smoothing the transition into higher education: What can be learned from student non-completion. *Journal of Institutional research*, *9*(1), 35-47.

Zeegers, P. (2001). Student learning in science: A longitudinal study. *British Journal of Educational Psychology 71*(1), 115-132.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary educational psychology*, *25*(1), 82-91.

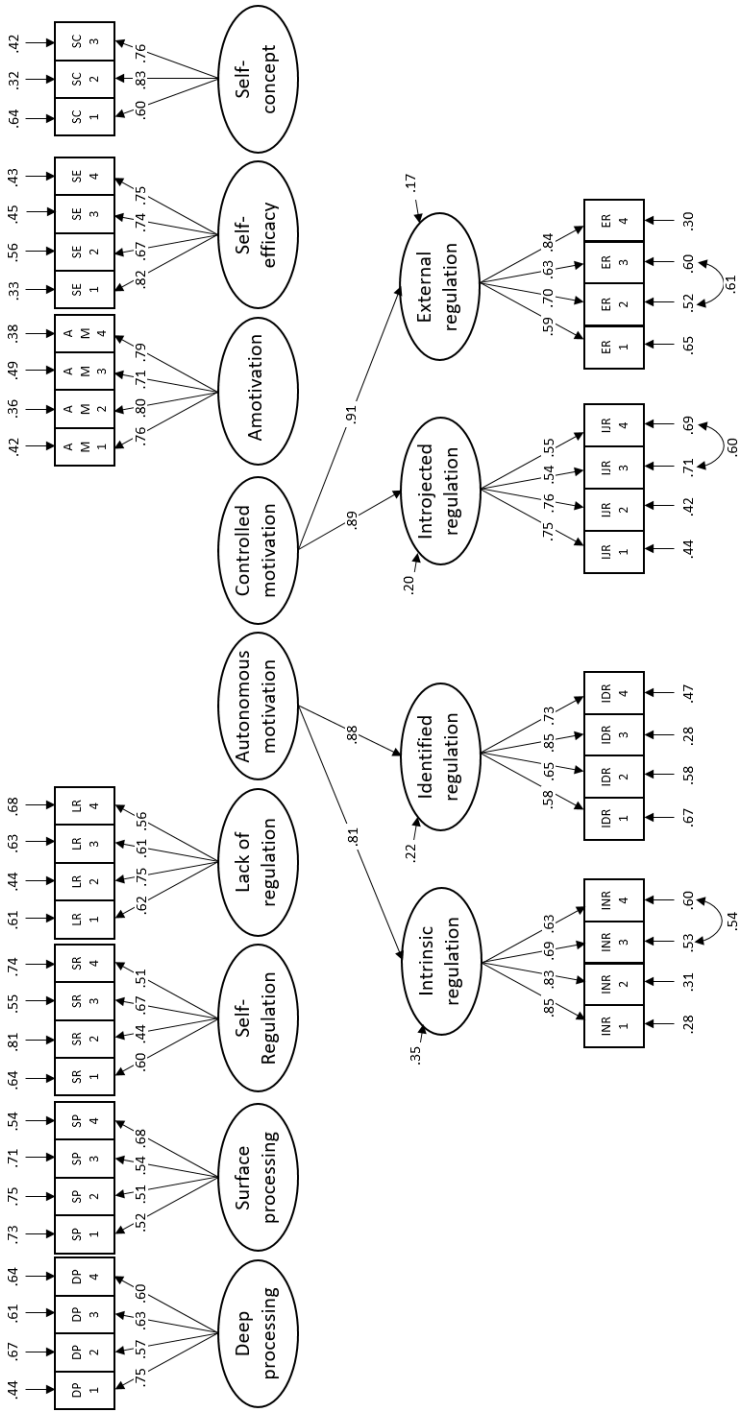
Zusho, A. (2017). Toward an integrated model of student learning in the college classroom. *Educational Psychology Review*, *29*(2), 301-324.

List of Figures

List of Tables

APPENDIX A

The following figure depicts the factor structure and details the standardized parameter estimates of the final CFA model. In this model we allowed the nine latent non-cognitive constructs under scrutiny to covariate. However, for the purpose of clarity, the estimates of these correlations were omitted from the figure and displayed in the table below. The figure further shows that the autonomous and controlled motivation factors are in fact second order latent variables. Their underlying first order factors are the four-item latent variables intrinsic regulation, identified regulation, introjected regulation and external regulation. The reader is referred to Deci and Ryan (2000) for more information and the theoretical rationale of this factor structure.



Parameter estimates of the final CFA model

Correlations between latent non-cognitive variables

		1.	2.	3.	4.	5.	6.	7.	8.	9.
1.	Deep processing	1								
2.	Surface processing	.18**	1							
3.	Self-regulation	.60**	.21**	1						
4.	Lack of regulation	-.40**	-.04	-.05	1					
5.	Autonomous motivation	.64**	.19**	.52**	-.25**	1				
6.	Controlled motivation	-.26**	.09	-.07	.49**	-.18**	1			
7.	Amotivation	-.40**	-.09	-.26**	.50**	-.64**	.50**	1		
8.	Self-efficacy	.49**	.25**	.22**	-.81**	.41**	-.41**	-.44**	1	
9.	Self-concept	.43**	.03	.15*	-.63**	.31**	-.37**	-.40**	.63**	1

**p<0.01, *p<0.05.

APPENDIX B

The incremental value of non-cognitive variables in the prediction of early academic achievement is the prime focus of the present study. Therefore, in the following paragraph, we provide more detailed information on the direct relationships between the non-cognitive determinants and the outcome measure, and on the final steps in the process of choosing the best-fitting model.

As stated in the manuscript, in a first step, we estimated the effect of each predictor variable (control, cognitive and non-cognitive variable) on the early academic achievement outcome separately, using correlation analysis (see Table 3 in the manuscript). Four non-cognitive scales did not have a significant direct effect on this outcome measure (autonomous motivation, controlled motivation, self-regulation and surface processing). With the five non-cognitive variables that did have a direct effect on early academic achievement (self-efficacy, self-concept, amotivation, deep processing and lack of regulation), we subsequently estimated a structural model that contains direct relationships of these variables in combination with control and cognitive determinants on early academic achievement. This model also incorporated all mediating (indirect) effects as depicted in Figure 2 in the manuscript. The results (see Table below) indicate that, for the five non-cognitive factors (within one integrated model), none of the direct effects on early academic achievement are significant.

Standardized parameter estimates for model containing direct relationships between early academic achievement and cognitive variables, control variables, and all non-cognitive variables that (in prior analyses) showed significant correlations with this achievement measure.

	Estimate (β)	***P* - value**
Prior Knowledge - Math	0.160	0.000
Prior Knowledge - Chemistry	0.272	0.000
Self-efficacy	0.053	0.385
Self-concept	-0.011	0.819
Amotivation	-0.047	0.256
Deep processing	0.017	0.733
Lack of regulation	-0.053	0.479
Age	-0.168	0.000
Gender	-0.105	0.003
Prior education - Math	0.130	0.000
Prior education - Physics	0.058	0.098

Fit-indices: CFI=0.90, RMSEA=0.04, SRMR=0.06, AIC=70722.889

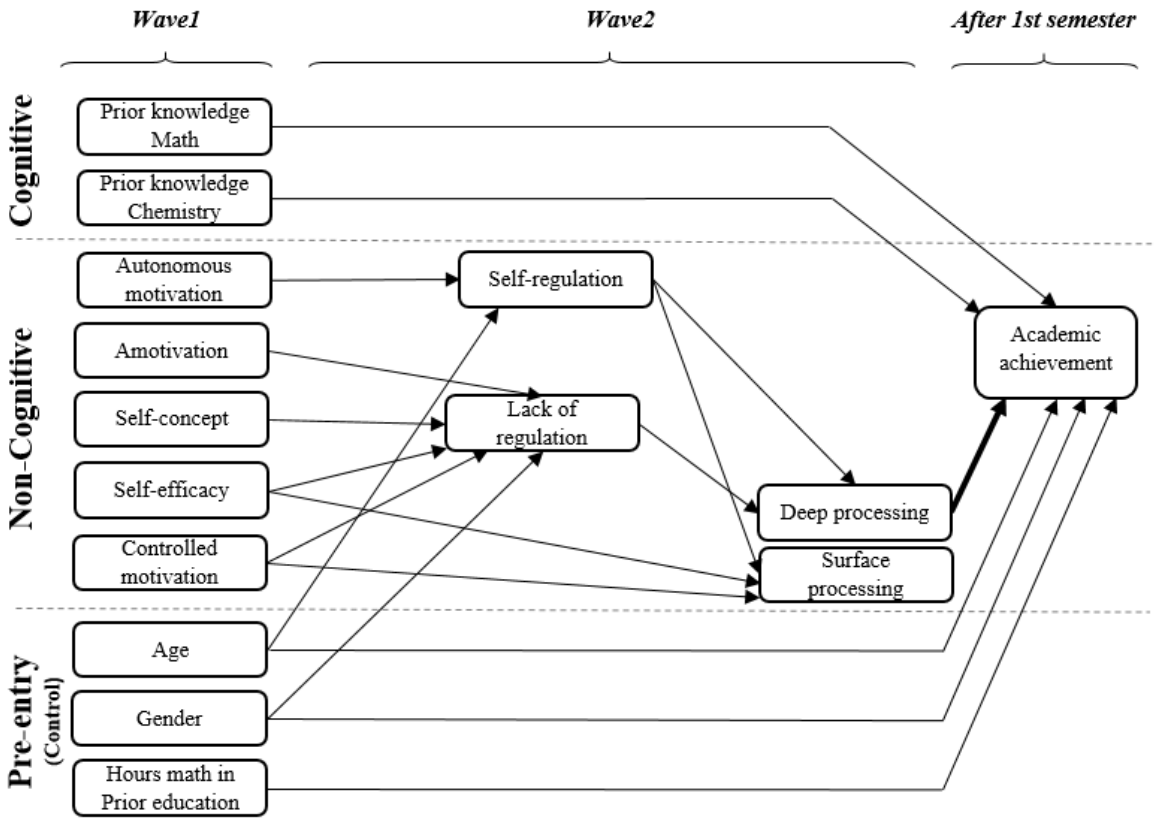
This absence of significant effects may be due to the fact that these non-cognitive factors are competing with one another for a small portion of unique variation in the dependent variable, and are related to each other, which may lead to redundancy effects (Cohen et al. 2013, Grewal et al. 2004). Hence, in line with the conceptual model (see Figure 1 in the manuscript), for each of the five non-cognitive predictors, the direct effect on early academic achievement was further explored. This resulted in five alternative models, with every alternative containing one direct significant relationship between academic achievement and one of the abovementioned non-cognitive determinants. Again, in each of this five

models, all mediating effects as depicted in Figure 2 (see manuscript) were included (by way of illustration, the model containing the significant direct effect of deep processing is shown in the figure below) .

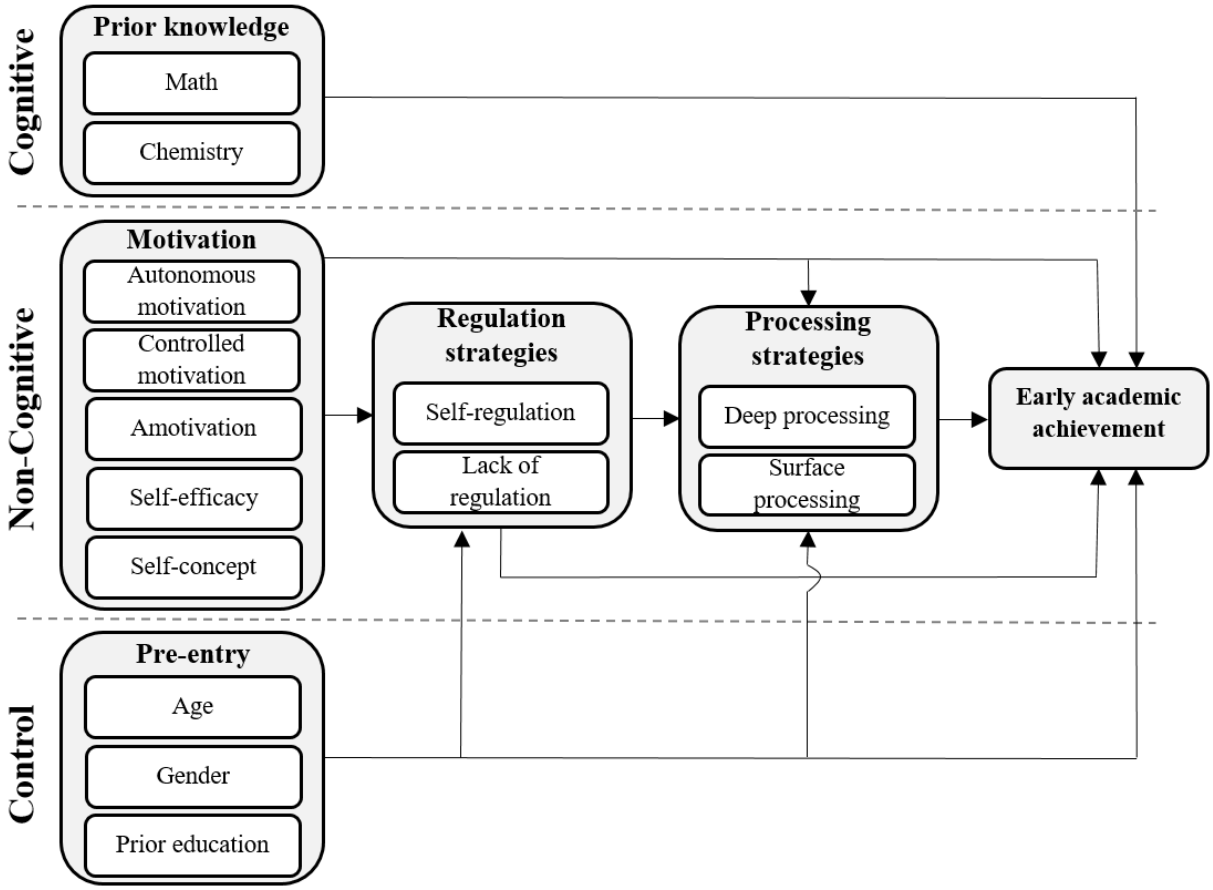
Fit indices and explained variance are shown in the following table. Based upon these results, we decided that the model containing lack of regulation as direct and significant predictor of early academic achievement was best-fitting. Therefore, this latter model, which is depicted in Figure 2 in the manuscript, was retained. The results of this model are described in the manuscript.

Overview of fit-indices and explained variance (R^2) of models containing different direct non-cognitive predictors of achievement

Model*	**CFI**	**RMSEA**	**SRMR**	**AIC**	**Chi-square**	**R^2**
Self-efficacy	0.905	0.037	0.063	70310.912	2242.1	0.275
Self-concept	0.904	0.037	0.063	70316.212	2247.4	0.276
Amotivation	0.904	0.037	0.063	70314.133	2245.3	0.280
Deep processing	0.904	0.037	0.063	70315.655	2246.9	0.276
Lack of regulation	0.905	0.037	0.063	70310.011	2241.2	0.280



Relationships between predictors and early academic achievement: model with direct effect of deep processing on academic achievement.


Cognitive
Prior knowledge
Math
Chemistry
Non-Cognitive
Motivation
Autonomous motivation
Controlled motivation
Amotivation
Self-efficacy
Self-concept
Regulation strategies
Self-regulation
Lack of regulation
Processing strategies
Deep processing
Surface processing
Early academic achievement
Control
Pre-entry
Age
Gender
Prior education

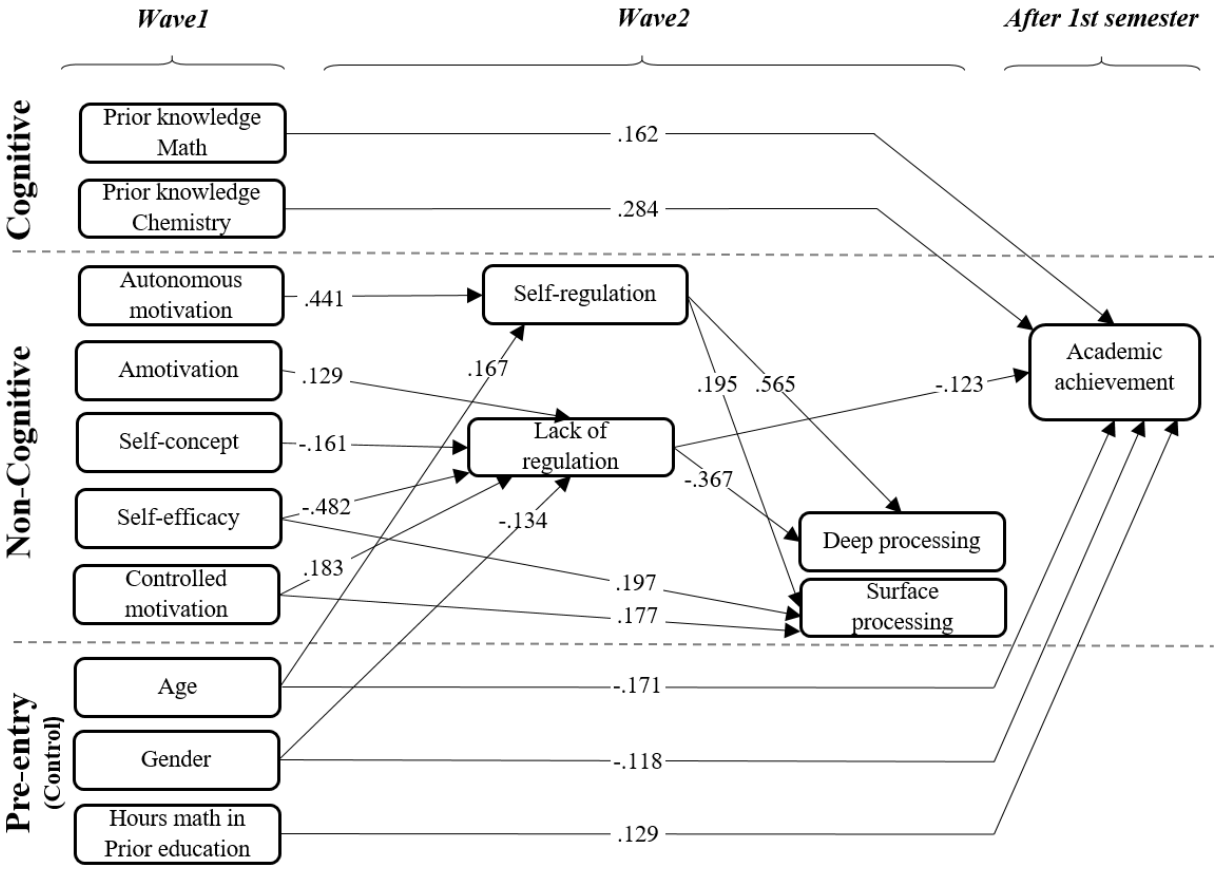

Wave1
Wave2
After 1st semester
Cognitive
Non-Cognitive
Pre-entry
(Control)
Prior knowledge Math
Prior knowledge Chemistry
Autonomous motivation
Amotivation
Self-concept
Self-efficacy
Controlled motivation
Age
Gender
Hours math in Prior education
Self-regulation
Lack of regulation
Deep processing
Surface processing
Academic achievement
.162
.284
.441
.129
-.161
-.482
.183
.167
-.134
.195
.565
-.367
-.123
.197
.177
-.171
-.118
.129

Table 1 *Hypothesised directions of the relationships under study*

	Self-regulation	Lack of regulation	Deep processing	Surface processing	Early academic achievement
Autonomous motivation	+	-	+	-	+
Controlled motivation	-	+	-	+	-
Amotivation	-	+	-	+	-
Self-efficacy	+	-	+	-	+
Self-concept	+	-	+	-	+
Self-regulation			+	-	+
Lack of regulation			-	+	-
Deep processing					+
Surface processing					-
Prior knowledge					+

Table 2 *Non-cognitive scales, measurement moment, number of items, item example and reliability*

Scale	Items	Wave	Example	α
Academic motivation [a]				
Controlled motivation	8	1	I am motivated to study, because I am supposed to do this.	.85
Autonomous motivation	8	1	I am motivated to study, because I want to learn new things.	.86
Amotivation	4	1	I am motivated to study… honestly, I don't know; I feel like I'm wasting my time in university.	.84
Self-beliefs [b]				
Self-efficacy	4	1	I have confidence in the way in which I study.	.81
Self-concept	3	1	I think I am well prepared for this course.	.79
Learning strategies [c]				
Deep processing	4	2	I try to combine the subjects that are dealt with separately in a course into one whole.	.72
Surface processing	4	2	I memorise lists of characteristics of a certain phenomenon.	.64
Self-regulation	4	2	In addition to the compulsory subject matter, I read other books or texts that had to do with this subject matter.	.63
Lack of regulation	4	2	I notice that it is difficult for me to determine whether I have mastered the subject matter sufficiently.	.73

Notes. Answering categories: a. 1 (Not important at all)– 5 (Very important); b. 1 (Completely disagree)- 5 (Completely agree); c. 1 (I never or hardly ever do this)– 5 (I almost always do this).

Table 3 *Intercorrelations among constructs*

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
1. *Prior education*-Hours math	1															
2. *Prior education*-Hours chemistry	.07	1														
3. *Prior education*-Hours physics	.16**	.74**	1													
4. Age	-.23**	-.06	-.18**	1												
5. Gender	.14**	.05	.08*	-.02	1											
6. *Prior knowledge*-Chemistry	.16**	.15**	.18**	-.10*	.09*	1										
7. *Prior knowledge*-Math	.31**	.06	.16**	-.30**	.08*	.55**	1									
8. Self-concept	.16**	.10*	.14**	.03	.15**	.21**	.12**	1								
9. Self-efficacy	-.04	.02	.04	-.06	-.04	.16**	.11**	.44**	1							
10. Amotivation	.06	.02	.07	-.10**	.11**	-.09*	-.06	-.18**	-.24**	1						
11. Autonomous motivation	-.16**	.02	-.05	.23**	-.22**	.04	-.06	.20**	.36**	-.38**	1					
12. Controlled motivation	.06	.01	.04	-.10**	-.00	-.07	-.00	-.19**	-.24**	.31**	-.17**	1				
13. Lack of regulation	.02	-.03	-.01	-.06	-.11**	-.10*	-.07	-.40**	-.50**	.31**	-.15**	.33**	1			
14. Self-regulation	-.05	.01	-.03	.18**	-.07	.02	-.06	.06	.06	-.18**	.33**	-.04	-.02	1		
15. Surface processing	-.06	-.05	-.06	-.00	-.05	-.05	-.02	-.01	.12**	.03	.08	.06	-.02	.14**	1	
16. Deep processing	.01	-.03	-.03	.13**	.04	.18**	.08	.24**	.26**	-.25**	.32**	-.21**	-.26**	.43**	.11*	1
17. Early academic achievement	.25**	.07	.17**	-.27**	-.05	.41**	.41**	.13**	.17**	-.09*	-.01	-.02	-.08*	-.07	-.06	.09*

**p<0.01, *p<0.05.

Table 4 *Overview of explained variance (R^2) by models containing different predictors of achievement*

Model	CFI	RMSEA	SRMR	AIC	R^2
Pre-entry factors	.91	.04	.06	71359.194	0.115
Pre-entry + Cognitive factors	.90	.04	.06	70265.959	0.277
Pre-entry + Cognitive + Non-cognitive factors	.91	.04	.06	70258.119	0.280
Non-cognitive factors	.90	.04	.06	71427.383	0.028