To Match or not to Match?

Improving Student-Program Fit in Dutch Higher Education

Karlijn Soppe

To Match or not to Match?

Improving Student-Program Fit in Dutch Higher Education

Karlijn Soppe

This research was embedded in the Interuniversity Centre of Educational Sciences (ICO).

ico

https://doi.org/10.33540/1069
ISBN: 978-94-6361-705-5

To Match or not to Match?
Improving Student-Program Fit in Dutch Higher Education

Matchen of niet Matchen?
Het Verbeteren van Student-Opleidingsfit in het Nederlands Hoger Onderwijs

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 7 juli 2022 des ochtends te 10.15 uur

door

Karlijn Femke Bernadet Soppe
geboren op 7 januari 1992
te Hardenberg

Promotoren:
Prof. dr. L.D.N.V. Wijngaards
Prof. dr. I.G. Klugkist
Prof. dr. T. Wubbels

Beoordelingscommissie:
Prof. Dr. S.F. Te Pas
Prof. Dr. A.G.J. Van de Schoot
Prof. Dr. J.W.F. Van Tartwijk
Dr. L. Hornstra
Prof. Dr. I.J.M. Arnold

TABLE OF CONTENTS

Preface 9

1 Introduction 11

1.1 Higher Education Dropout 13

1.1 The Dutch Context 13

1.2 The Implementation of Matching Procedures in the Netherlands 14

1.3 Person-Environment Fit 16

1.4 Research Design 17

1.5 The Present Dissertation: Evaluation of Matching Procedures 18

2 Do They Match? Prospective Students' Experiences with Choosing University Programs 23

2.1 Introduction 25

2.2 Program Choice in Higher Education in the Netherlands 27

2.3 Aim of the Study 28

2.4 Methods 29

2.4.1 Sampling & Participants 29

2.4.2 Ethical Approval 30

2.4.3 Procedures 30

2.4.4 Analysis 30

2.5 Results 31

2.5.1 Statements Confirming the Possibility to Test Person-Environment Fit 33

2.5.2 Statements Disconfirming the Possibility to Test Person-Environment Fit 34

2.5.3 Contradictory Statements Regarding the Possibility to Test Person-Environment Fit 35

2.5.4 Impact of Matching on Program Choice 35

2.5.5 Choice Certainty 39

2.6 Conclusion and Discussion 39

3 Determining Fit: The Role of Matching Procedures in Prospective Higher Education Students' Enrolment Behavior 45

3.1 Introduction 47

3.2 Person-Environment Fit 47

3.3 Matching Procedures in the Netherlands 48

3.4 Methods 51

3.4.1 Sample 51

3.4.2 Procedure 52

3.4.3 Analyses 52

3.5 Results 53
3.6 Discussion 59

4 Pre-enrolment Predictors of First-Year Academic Success: Indicators of Student-Program Fit for Different Disciplines 63
4.1 Introduction 65
4.2 Pre-Enrolment Predictors of Academic Success 66
4.2.1 Conscientiousness 68
4.2.2 High School Performance 68
4.2.3 Ability Beliefs 69
4.2.4 Interests 70
4.2.5 Program Orientation 70
4.2.6 Age and Gender 71
4.3 Differential Prediction by Discipline 71
4.4 Research Questions and Hypotheses 72
4.5 Study 1 73
4.5.1 Method 73
4.5.2 Analyses 74
4.6 Results 75
4.6.1 Descriptive Statistics 75
4.6.2 Structural Model 76
4.6.3 Differences by Discipline 77
4.7 Study 2 79
4.7.1 Method 79
4.7.2 Analyses 81
4.8 Results 81
4.8.1 Descriptive Statistics 81
4.8.2 Structural Model 82
4.8.3 Differences by Discipline 82
4.9 General Discussion 83
4.9.1 Limitations and Directions for Future Research 85
4.9.2 Implications for Research and Practice 85
4.10 Conclusion 86

5 Pre-university Motivation and First-Year Study Success: Text Mining of Pre-enrolment Questionnaires 89
5.1 Introduction 91
5.2 Methods 93
5.2.1 Sample 93
5.2.2 Measures 93

5.2.3 Preprocessing the Motivation Statements 95
5.2.4 Feature Engineering 96
5.2.5 Training the Algorithm 97
5.2.6 Analysis 97
5.3 Results 99
5.3.1 Model 1: Student Characteristics 100
5.3.2 Model 2: Only Text 100
5.3.3 Model 3: Text and Text Features 100
5.3.4 Model 4: Student characteristics and text 102
5.3.5 Model 5: Student Characteristics and Text Features 103
5.3.6 Model 6: Student Characteristics, Text, and Text Features 104
5.3.7 Comparing Most Frequently Used Terms of Correctly and Incorrectly Classified Dropouts 105
5.4 Discussion 105

6 Discussion 111
6.1 Introduction 113
6.2 Summary of the Main Findings 113
6.3 Effectiveness of Matching Procedures 115
6.3.1 Student Perceptions of Usefulness 115
6.3.2 Making Students at Risk of Dropout Reconsider Their Program Choice 116
6.3.3 Improving First-Year Academic Success 117
6.4 Scientific Contributions 118
6.5 Practical Implications 119
6.6 Limitations and Future Research 122
6.7 Conclusion 124

References 125
Appendices 139
Nederlandse Samenvatting 143
About the Author 151
ICO Dissertation Series 153
Dankwoord 157

PREFACE

Like many prospective students, when I had to choose what I wanted to study, I struggled. I think I changed my mind at least 5 times, before eventually landing on my final choice: Sociology. In high school, I had opted for a more science-oriented set of subjects to keep my options for university open. Being in a more science-oriented school track made me initially explore science-oriented programs like Artificial Intelligence or Human Movement Sciences, as well as programs at universities of applied sciences like Physiotherapy and Psychomotor Therapy. However, science subjects were not my strong suit. Languages were more my thing, so I also explored the option of studying German Language and Culture. None of the programs really convinced me though. Was I really able to pull of physics at the university level? Was I interested enough in German to study it for several years? I was a little lost. That is, until I learned about the concept of social sciences. I visited an open day of a Sociology program and while listening to a lecture about adolescent criminal behavior, I knew that this would be something to explore further.

Once I decided to study Sociology, I visited “student-for-a-day” activities at two different universities. At both programs I really felt at home, I liked the students and staff and enjoyed the lectures as well as the cities. Eventually, I enrolled in the bachelor program at Utrecht University. Was I certain of my choice? Absolutely. Was I well-informed? Maybe not so much. Despite my extensive orientation, I quickly learned that the program consisted of more statistics courses than I had anticipated and that the vast amounts of literature I had to cover on a weekly basis, exceeded my reading skills. I felt a little lost again. Had I chosen the wrong program after all? Did I lack the required skills for the program of my choice?

Today, at the end of my PhD I know that the transitional phase from high school to university is generally considered difficult and that many students struggle in the first few months of their university-experience. I have read about it in scientific literature, witnessed it in students that I taught, and I had experienced it myself when I started studying. Maybe my personal experience is why I was so drawn to this PhD project. It is definitely one of the reasons that kept me motivated throughout this journey.

Chapter 1

Introduction

1.1 Higher Education Dropout

Student dropout is one of the biggest challenges in higher education. Dropout is associated with negative consequences for both higher education institutions and individual students. Students might start questioning their capabilities and experience financial consequences as a result of dropout. Universities are often partially funded based on graduation rates (Jongbloed et al., 2018; Kirk, 2018), so low retention rates result in low returns on their investments.

One of the core reasons of dropout worldwide is a wrong initial program choice (e.g., Bean, 2005; O'Keefe et al., 2010; Willcoxson & Wynder, 2010; Yorke, 2000). In many educational systems, other core reasons include financial problems or conflicts with work and/or family commitments. However, in the Netherlands, a wrong program choice has been the most important reason for dropout reported by students over the last decade (Van den Broek et al., 2020). Students who do not match with the program are more likely to drop out than students who do experience program-fit (Feldman et al., 1999). A fitting program choice is associated with positive outcomes for the student. Students can assess their fit with a program on several aspects. For example, experiencing a sense of belonging positively influences students' engagement in the academic process (McFarlane, 2018; Kirk, 2018) and is associated with a lower probability of dropout (Trowler, 2010). On the other hand, a lack of fit between student and program is more likely to result in the student dropping out. Previous work highlights that experiencing feelings of misfit in general (Feldman et al., 1999; Ulriksen et al., 2010; Warps et al., 2017) or feeling a lack of sense of belonging more specifically (Naylor et al., 2018; Tinto, 1987) are among the most important predictors of dropout. These feelings of misfit become clear when students, once they started studying, realize that their expectations do not fit with the reality of the program (e.g., Warps et al., 2017, p.11), which could result in dropout once more realistic beliefs set in (Watson et al., 2004). In fact, having unfulfilled expectations is the second most important reason of dropout in the Netherlands in the last decade (Van den Broek et al., 2020).

1.2 The Dutch Context

The educational system in the Netherlands is highly stratified. Around the age of 12, after eight years of primary school, children are sorted into specific tracks of secondary education. This track is determined based on the advice of the teacher and a school-leaving test, administered in the final year of primary school. The track of secondary education, in turn, determines which type of tertiary education students can enroll in. Tertiary education in the Netherlands consists of three levels, of which the highest two are considered higher education elsewhere in the world.

Dutch higher education consists of universities of applied sciences (Dutch: hbo) and research universities (Dutch: wo). Only students with a secondary diploma at the "preparatory university" level (Dutch: vwo), can directly enroll in a research university. Roughly 20% of all students follow this 6-year preparatory university education (Nederlands Jeugdinstituut, 2021). A second route to research universities is by obtaining a propaedeutic diploma or bachelor's degree at a university of applied sciences. The high stratification in Dutch education results in a relatively homogenous group of applicants to higher education regarding cognitive abilities.

In the Netherlands, selective admission procedures are limited to certain bachelor programs, mainly in the medical field (e.g., Medicine or Pharmacy), some technical programs (e.g., Aerospace Engineering or Computer Science and Engineering), business studies (e.g., International Business Administration or related programs), and Psychology. All other programs are open-admission programs, sometimes requiring a certain high school track (e.g., a high school track with emphasis on science to study Physics). All things considered, prospective students in the Netherlands have a wide range of options for their higher education program choice.

1.3 The Implementation of Matching Procedures in the Netherlands

To increase academic success in higher education, the problem of low retention was put on the political agenda in the Netherlands. In 2013, the Dutch government passed a law (Wet Kwaliteit in Verscheidenheid Hoger Onderwijs [Quality in Diversity Law] 2013) that, since the academic year 2014/15, among other things, gives prospective students the right to have a final check on their program choice before starting higher education. Because there is no selection for most programs in the Netherlands, the parliament deemed the study check important for getting every prospective student at the right place. The process for the final check is called matching and is offered after a prospective student has filed an initial admission request.

When implementing the matching procedures, it was assumed that the decision process of prospective students would go as follows: prospective students will orientate themselves on the choice of study by means of open days and potentially participation in trial studying activities, make a choice and then file an initial admission request for a study program. Then students will participate in a matching procedure where they are either directly advised or are encouraged to self-reflect on their program choice. It was assumed that most students would be rightly confirmed in their choice, after participating in the matching procedure. Amongst the (small) group of students who have doubts, several will decide not to start the program. Another part of the students, now better prepared by the matching, will still start the program. The assumption that fewer students make the wrong program choice and that all students start their study better prepared as a result of matching will lead to higher returns and less drop-out in the first year.

All universities agree on the aim of matching, getting the right student at the right place (VSNU, 2017). They also agree that this can be achieved by giving prospective students a realistic view of the content, level, teaching methods, and job market of a program. Moreover, they all acknowledge the need to build a sense of belonging with the program and institution prior to enrolment. Lastly, Dutch universities agree that a good fit between student and program can be achieved if program staff can help students reflect on their program choice through insight in their expectations, motivation and interests.

The way in which different universities are trying to achieve these goals differs, with a variety of matching types as a result. The implementation of matching does not only vary between universities but may also vary between programs within the same university. Most matching procedures start with an online questionnaire, which generally contains sections on the following: high school grades; orientation activities undertaken to learn about different programs (e.g., visiting an open day, browsing the website); motivation / reasons for choosing the program; expectations of the program; expected future jobs; ability beliefs; general time use; and expected number of contact hours per week. Sometimes concepts regarding personality (e.g., conscientiousness and/or openness to experiences) or learning styles are included as well. Many universities offer a follow-up activity after prospective students have completed the questionnaire. The most common activities are so-called matching days on campus and online courses (sometimes combined with homework and/or a test). A third activity that is frequently part of matching procedures concerns interviews with program staff. These interviews are generally only held with students who are deemed at risk of dropout based on their answers on the questionnaire. In some universities these follow-up activities (i.e., matching day, online course, or interview) are compulsory, either for all students or only for students deemed at-risk of dropout. Matching procedures are always concluded with some type of feedback. In many universities it has taken the shape of a concrete advice, often in terms of a traffic light analogy (i.e., green advice for a good fit, orange for doubt and red advice for a deemed misfit). A minority of the universities provide more generic feedback (based on questionnaire and/or activity) in an attempt to spark self-reflection.

1.4 Person-Environment Fit

As mentioned in the beginning of this introduction, a wrong program choice is the most important reason of higher education dropout in the Netherlands. Thus, if one wants to reduce dropout in the Netherlands, the focus should lie on preventing a wrong program choice and thus, universities have designed their matching procedures with this in mind. Throughout this dissertation we argue that matching procedures are likely to foster a "good" program choice if they allow students to test whether they match with the program.

The assumption that outcomes are a function of the interaction between individuals and their environments stems from person-environment interaction theory (Lewin, 1935). Person-environment fit is defined as the compatibility between individual and environmental characteristics (Kristof-Brown et al., 2011). Fit research across a variety of domains shows that an individual's performance improves if there is alignment between a person and their environment (Ward & Brennan, 2020). Within the educational context, person-environment fit builds on the assumption that students with certain characteristics are more likely to choose certain programs (Astin, 1993) and that congruence between student and program is paramount to academic success (Feldman et al., 1999). We argue that students who can test their fit with the program of their choice prior to enrolment will make a better choice.

Which aspects of fit are important for a higher education choice is not clear and there is little research on this topic. For this dissertation we integrate and build on several theories, of which two motivational theories, Expectancy-Value Theory (EVT; Eccles & Wigfield, 2002) and Self-Determination Theory (SDT; Deci & Ryan, 1985), as well as Tinto's Student Integration Model (1977) are the most important. We identify three concepts that we assume to be important in the context of a higher education program choice. Each of these concepts will be explained below.

First, we consider the concept ability beliefs. This concept is defined in this dissertation as "one's beliefs in their abilities to perform a certain task". In the literature it is referred to as self-efficacy (Bandura, 1977), beliefs about competence (Eccles-Parsons et al., 1983) or ability beliefs (Eccles et al. 1989). In Expectancy-Value theory, it is proposed that someone's beliefs about their competence influences their expectancies and values. In return, these expectancies and values directly influence performance and task choice. Thus, in the context of a higher education program choice, prospective students will assess their current capabilities and their estimated probability of success in order to determine fit on this aspect. A closely related concept is competence, one of the three basic psychological needs, identified by Ryan and Deci

(2017) in their Basic Psychological Needs Theory, a sub-theory of SDT. Competence concerns experiencing mastery of a specific skill (Vansteenkiste et al., 2020). In the educational context, this basic need becomes satisfied when students successfully engage in a certain task and experience opportunities to extent their skills regarding this task. Based on insights from these three theoretical perspectives, we assume that the concept of ability beliefs is a vital aspect to experience student-program fit.

Second, vocational interest plays an important role in adolescents' learning and development (Renninger & Hidi, 2017). Vocational interest is related to program choice (Whitney, 1969), but it might be difficult for students to decide which interests to pursue. Students often have multiple interests, which they can generally not pursue all within one higher education program (Hofer, 2010; Vulperhorst et al., 2018). According to EVT, students will weigh the value of the program (e.g., for example in terms of their different interests), thereby trying to maximize pros and minimize cons (Eccles & Wigfield, 2002) in their program choice. Moreover, Deci & Ryan (2000, p.57) "*recognize that basic need satisfaction accrues in part from engaging in interesting activities*". In this dissertation, interests are defined as "the extent to which a person values certain topics over others". We assume that students need to acquire a realistic sense of how their interests align with the program they wish to pursue and vice versa.

Third, we argue that students need to feel a sense of belonging in order to stay motivated. SDT identifies relatedness as another basic psychological need (Ryan & Deci, 2017). Relatedness indicates feelings of warmth and bonding (Vansteenkiste et al., 2020). This basic need becomes satisfied when a person feels connected with and important to others. In his Student Integration Model, Tinto (1975, 1993) identifies the need for social integration as an important factor to prevent student dropout. Tinto defines social integration as the presence of positive relationships with peers. We use a somewhat wider scope, by defining sense of belonging as "a sense of connectedness with fellow students, staff members and one's physical surroundings". In this dissertation, sense of belonging is the third aspect that we deem important for students in determining fit with the program of their choice.

1.5 Research Design

The studies presented in this dissertation are based on data from several Dutch research universities. Since information on matching procedures is only available within the institutions themselves, the contribution of these universities is vital for the comparison of the different types of matching procedures described in this dissertation. At the start of this PhD-trajectory, based on types of matching procedures and the presence of specific programs, we selected seven universities we wished to include in this research. Eventually, four universities agreed to participate in both the qualitative and quantitative research components of this dissertation. Within these universities, we chose four programs that were to a certain extent representative of a) a typical humanities program, b) a typical STEM (Science, Technology, Engineering, Mathematics) program, c) a typical social sciences program, and d) a program that is more frequently than usual chosen by students who do not know what they want to study.

These programs offered a variety of different matching procedures, which made comparisons between types of matching possible. The aim was to compare different types of procedures, rather than different institutions, but there is strong overlap between the type of matching procedures and the universities offering these types. To maintain confidentiality, universities and programs will be addressed as numbers and/or letters throughout the separate studies in this dissertation.

1.6 The Present Dissertation: Evaluation of Matching Procedures

Matching procedures have been in place for a while now and in this dissertation the focus lies on researching their effectiveness at Dutch research universities. The central question in this dissertation is *whether and to what extent various types of matching procedures at Dutch research universities are effective*. To answer that question, effectiveness needs to be defined. In this dissertation, the effectiveness of matching procedures is defined by considering the goal of matching: getting the right student in the right place. Moreover, although it is not explicitly defined in any policy report or parliamentary document, a subsequent goal of matching is to increase academicsuccess, specifically first-year retention and time to degree. With these two goals in mind, in this dissertation an effective matching procedure is operationalized as a procedure that 1) is considered useful by students in their final program choice, 2) makes students who are deemed at-risk of dropout reconsider finalizing their enrolment, and 3) is associated with first-year academic success. The present dissertation examines how different elements of various types of matching procedures are associated with student enrolment behavior and academic outcomes (see Figure 1.1), using different types of research methods. The figure shows the transition into university and the first year of studying on the top row. At any point, students can drop out of the enrolment process or the program (vertical lines). Upon dropout, they will have to make a new program choice and possibly participate in a new matching procedure.

In Chapter 2 we ask ourselves *how prospective students perceive the role of matching in their program choice and how they perceive each of the elements of matching in testing the components of person-environment fit*. To answer these questions, we conducted 61 interviews with prospective students of four Dutch universities. Interviewed students have participated in different types of matching procedures, which allows for comparison of different elements of matching and how students experience their role in program choice. The data on student experiences are analyzed in NVivo, using a grounded theory approach. As a result of the coding process, we construct a conceptual model on person-environment fit for the matching context.

Chapter 3 provides insight in the association between different types of matching procedures and student enrolment behavior. We use the conceptual model from Chapter 2 as our theoretical lens to answer the questions *how enrolment rates vary between university programs with different types of matching procedures and how these rates differ before and after implementing matching procedures*. Differences in enrolment rates of thirteen programs at four Dutch universities are visualized and tested in SPSS. By studying enrolment rates across programs with different types of matching procedures as well as over time, we assess the effectiveness of these matching procedures for enabling prospective students to make an informed program choice.

In Chapter 4 we investigate the relation between pre-enrolment indicators of person-environment fit and first-year academic success, using Structural Equation Modelling (SEM) in MPlus. This study evolves around several research questions. First, we wonder *whether indicators of fit, as measured prior to enrolment, can predict first-year GPA and earned credits*. Second, *we explore whether the relation between the pre-enrolment indicators of fit and academic success differs between disciplines*. In this study we first test our SEM model on the data of one university. Then we replicate these findings, using data from another university. Hence, our last research question in this study concerns *the extent to which our hypothesized model yields similar results across universities.*

In Chapter 5 we perform text mining techniques in Python to dive deeper into students' written motivation. The questions we aim to answer in this study are *whether students at risk of dropout can be identified through text mining, based on their motivation for the program of their initial choice as written in their intake questionnaire prior to enrolment; and whether information extracted from these motivation statements add predictive power net of student characteristics.* By analyzing short motivation statements from the intake questionnaires of one university by means of various Natural Language Processing techniques, we assess whether motivation, as analyzed through text mining, can predict first-year dropout. Moreover, by comparing and combining the text data with the pre-enrolment indicators of fit from the study presented in Chapter 4, we determine whether analyzing text data can provide us with additional predictive power for research on student dropout.

Chapter 6 summarizes and discusses the main outcomes from the studies presented in Chapters 2-5. Furthermore, we address the methodological limitations, discuss the implications on guiding students in the transition into higher education and their performance during the first year, and provide suggestions for future research. Lastly, based on the research in this dissertation, recommendations are made for matching procedures in Dutch higher education practice.

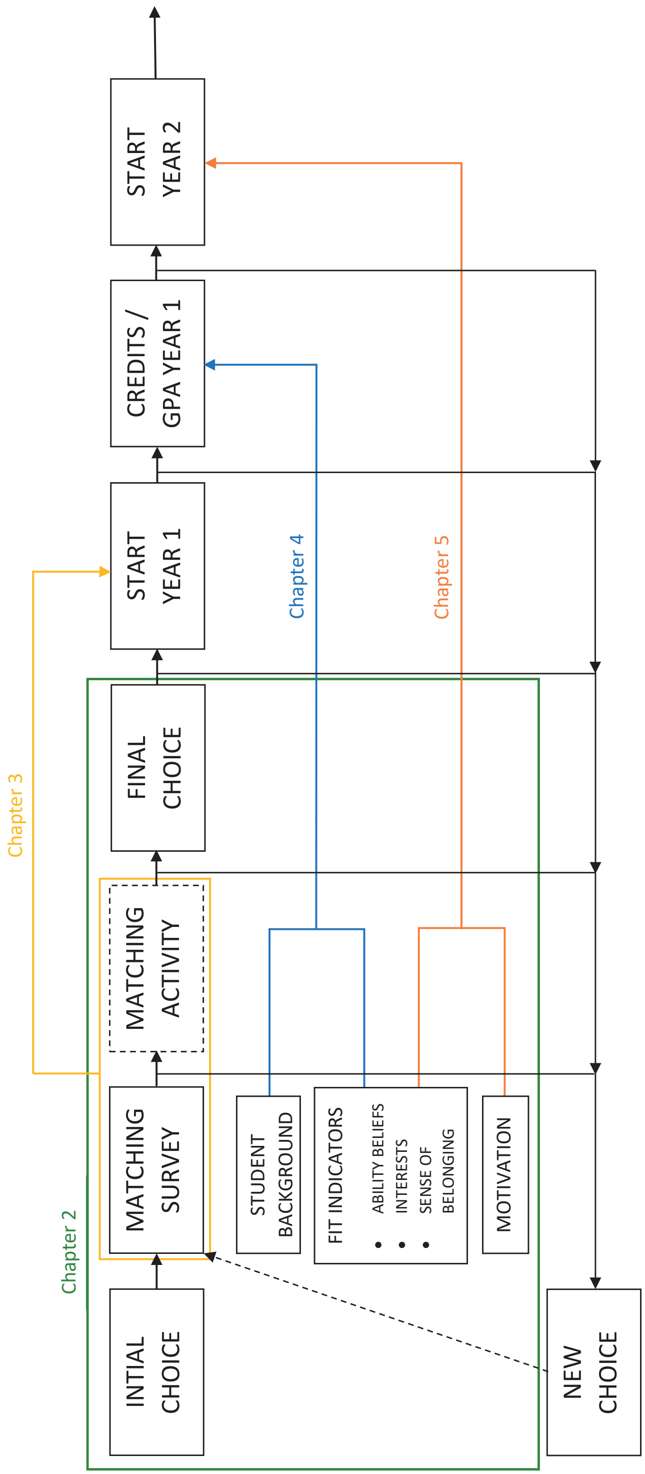

Figure 1.1 Schematic overview of the concepts and relations researched in this dissertation.

Chapter 2

Do They Match? Prospective Students' Experiences with Choosing University Programs

This chapter is published open access as: Soppe, K. F. B., Wubbels, T., Leplaa, H. J., Klugkist, I. G., & Wijngaards-de Meij, L. D. N. V. (2019). Do they match? Prospective students' experiences with choosing university programmes. *European Journal of Higher Education, 9*(4), 359-376.

ABSTRACT

When transitioning from high school to university, young people must choose a programme that fits them. We argue that prospective students who can test this fit before starting the programme, will make a better choice. We propose an integrated framework where testing person-environment fit on ability beliefs, interests and sense of belonging possibly contributes to making the right choice. Dutch matching procedures are supposed to serve as a fit-test for prospective students choosing a university programme. 61 prospective students at four Dutch universities were interviewed on the role of matching in their programme choice. Different elements of matching appear to allow for testing fit but vary in which aspects of fit can be tested and the impact they have. It can be cautiously stated that the more aspects of fit that can be tested, the more a matching procedure impacts prospective students' final programme choice.

Author contributions: KS, LWM, TW, & IK designed the study. KS & LWM recruited partner universities. KS organized the data collection, conducted the interviews, and conducted initial coding. KS & HL analyzed the data and built the conceptual model. KS wrote the paper. TW, HL, IK, and LWM provided extensive feedback on the manuscript.

2.1 Introduction

When transitioning from high school to university, young people must choose a program that fits them. The difficulty of this task is particularly reflected in academic success rates. Since several decades academic success in higher education is an important topic in higher education research. This is evident from the vast amount of scholarly attention in the UK (e.g., Johnes & McNabb, 2004), USA (e.g., Betts & Morell, 1999; Rausch & Hamilton, 2006), Australia (e.g., De Rome & Lewin, 1984), Russia (e.g., Tolstova, 2006), and many European countries (EMBO, 2006). Moreover, student retention or dropout is very high on the policy agenda in about three quarters of the European countries surveyed recently by order of the European Commission (Vossensteyn et al., 2015). Dropout is an important issue, because it has various negative consequences for both the student and the university. Students may start questioning their competences due to feelings of failure after dropping out of the program of their initial choice. Moreover, low retention rates in higher education are associated with financial consequences for both the university and the student, because university funding is often partially based on the number of students who graduate without a certain period of delay. Although the percentage of university students who graduate may be approximately 90% (e.g., in Japan), it can be as low as approximately 50% (e.g., in Italy) (European Commission, 2015).

Although additional coincidental factors might contribute to the student being successful during the university program, we argue that person-environment fit, which can be defined as the compatibility between individual and environmental characteristics (Kristof-Brown & Guay, 2011), is important in choosing a program in which a student is likely to succeed. The assumption, that outcomes are a function of the interaction between individuals and their environments, stems from person-environment interaction theory (Lewin, 1935). Within the educational context, person-environment fit builds on the assumption that students with certain characteristics are more likely to choose certain programs (Astin, 1993). Additionally, some research suggests that congruence between the student and the program is paramount to the academic success of college students (Feldman et al., 1999), i.e., students who lack feelings of fit are less likely to graduate. Hence, person-environment fit is important for program choice. We argue that prospective students who can test this person-environment fit on different aspects before making a final program choice will make a better choice. We now introduce aspects on which person-environment fit could be tested based on expectancy-value theory (EVT; Eccles & Wigfield, 2002) and self-determination theory (SDT; Deci & Ryan, 1985).

Two of the aspects on which prospective students may consciously or unconsciously assess their fit are ability beliefs and interests. According to EVT choosing to perform a certain achievement-related task can be explained by ability beliefs (i.e., how well I can perform on the task) (Bandura, 1977) and the value of the activity (e.g., interest in the course materials) (Atkinson, 1957; Wigfield, 1994; Wigfield & Eccles, 1992). In other words, prospective students will weigh the costs and benefits of a program by assessing the relative value of the program, their current capabilities to start the program and their estimated probability of success. Closely related to the ability beliefs concept is the competence concept in SDT. From SDT it is argued that among other things, feelings of competence result in intrinsic motivation (Deci & Ryan, 2008). Competence involves the understanding of how to acquire various outcomes and being effective in the essential actions to achieve these outcomes.

Furthermore, from SDT it is argued that relatedness, the social component in adjusting to a new environment, results in intrinsic motivation (Deci & Ryan, 2008). Developing social connections with others, or experiencing a sense of belonging (Freeman et al., 2007) can be comforting on both social and academic aspects and allows for enhanced ability to cope with the demands of the transition to higher education (Hoffman et al., 2002). Sense of belonging has often been investigated in the context of minorities, such as female students in Science, Technology, Engineering and Mathematics (STEM) programs (e.g., Ulriksen et al., 2010). However, it has been argued that experiencing a sense of belonging is vital for all students (Feldman et al., 1999; Strayhorn, 2012).

Lastly, SDT includes the concept autonomy, which refers to initiating and controlling one's own actions. According to this theory, greater feelings of autonomy enhance intrinsic motivation (Deci & Ryan, 2008). In numerous countries, studying at the university is accompanied by prospective students moving to live on their own in a new city. In that case the transition to higher education will be associated with an increase in (feelings of) autonomy. For example, by moving out of the parental home, prospective students will get many new responsibilities on a personal level (e.g., cooking and cleaning) and students' parents will no longer set their rules. We argue that this expansion of autonomy is overwhelming, and thus, prospective students may not take autonomy within the program into account when choosing a program.

We propose in Figure 2.1 an initial framework for testing person-environment fit combining the aspects from several theories, and stating that interests, abilities and sense of belonging are elements on which person-environment fit can be assessed in program choice in higher education. These elements do not have to be of equal importance to a student in making a choice, but they are all assumed to be important components that prospective students consider.

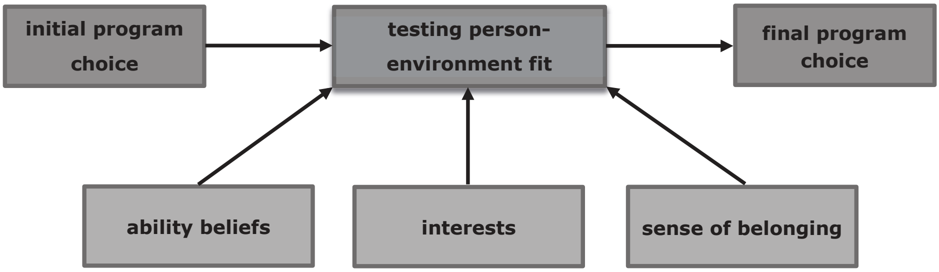

Figure 2.1 Assumed variables in choosing a program.

2.2 Program Choice in Higher Education in the Netherlands

With a completion rate in higher education of approximately 70%, the Netherlands holds an average position in dropout rates within Europe. In countries with admission procedures in higher education like those in the Netherlands, comparable figures are found (e.g., Belgium 73%; Denmark 81%; European Commission, 2015). Other statistics show that in the Netherlands, approximately 20% of university students drop out in their first year (VSNU, 2016). Of the students who finish their bachelor's program, only a quarter of the students finish within the normal three-year timeframe (Statistics Netherlands (CBS), 2017a) and about three-quarters earn a diploma within seven years (Statistics Netherlands (CBS), 2017b). The clear majority drops out in the first year of the program and this implies that assisting prospective students in their program choice is important.

Australian research has shown that prospective students expect university to be different than high school, but expectations of university life proved unrealistic (Crisp et al., 2009). Based on these and other insights in the transition process, some authors call for interventions to facilitate the transition into and adjustment to university (e.g., Brinkworth, et al., 2008; Mattanah et al., 2010). The Netherlands has implemented procedures that should allow prospective students to make more informed choices.

Since 2014 prospective students in the Netherlands have the right to have a final check on their program choice before starting higher education, to help them in the process of program choice. This final check is called matching and is offered after a prospective student requests admission. The implementation of matching varies between programs. Examples of possible activities are (online) questionnaires, lectures on campus, interviews, and online courses. A combination of several elements that may or may not be compulsory is the most common format of matching.

Several elements of matching are considered in this study. Every program first sends out an online questionnaire. Questionnaires are not the same across universities, but generally contain similar sections such as (a selection of): high school grades; attendance at orientation activities to learn about different programs; motivation / reasons for choosing the program; expectations of the program; expected future jobs; ability beliefs; general time use. This online questionnaire is often followed by a brief online course, a personal interview, or a matching day. In an online course, prospective students receive information (e.g., through online lectures) on a topic representative of the program, followed by a test. Personal interviews with program staff are often obligatory for prospective students who are deemed at risk of drop-out based on their answers on the questionnaire. Lastly, a matching day is a day on campus where prospective students attend lectures, tutorials or lab sessions that are aimed at giving them a representative study experience; generally, the prospective students must prepare for that day at home by completing assignments, e.g., based on reading materials.

Upon completion of the matching, program staff may provide prospective students with personal advice regarding their fit with a specific program, more generic feedback, or neither. Programs that provide advice state how well they think a prospective student fits into the program. Programs that provide more generic feedback provide prospective students with statements about their fit without adding advice, leaving it to the prospective students to assess how well they fit. At some universities, all activities are compulsory; and at others, some of the activities, or none, are compulsory.

2.3 Aim of the Study

Recently, research on program choice and admission to Higher Education mainly evolves around selective admission and (types of) selection procedures (e.g., Barrow et al., 2018; Niessen & Meijer, 2017; Steenman et al., 2014; Umarji et al., 2018). Program choice and admission in open-admission contexts, however, is an under-studied topic. We focus exactly on this aspect of admission: person-environment fit in the open-admission context of Dutch matching procedures. The constructed framework on person-environment fit provides a context in which the role of the various matching activities in making a program choice can be studied. To identify how matching contributes to program choice through assessing person-environment fit, the following questions will be addressed: "*How do prospective students perceive the role of matching in their program choice?*" and "*How do prospective students perceive each of the elements of matching in testing the components of person-environment fit?*"

2.4 Methods

In this qualitative study, data were collected by two interviewers, the main researcher and a research assistant, through semi-structured phone interviews from mid-April to the end of August 2017. The interviews were conducted as soon as possible after the matching and always before the start of the academic year.

2.4.1 Sampling & Participants

For this study, we found cooperation of seven university programs at five Dutch universities. These programs were chosen based on several criteria: diversity in matching activities, programs in the same discipline that were offered at several universities, and both STEM and non-STEM programs. Contact persons from the different partner universities decided whether to conduct a passive or active consent procedure with their prospective students. In both cases, prospective students received an email for the consent procedure, in which they were also informed of the possibility of winning one of five gift cards with a value of €25. We strived for approximately 10 interviews per program (i.e., a total of 70 interviews), to ensure enough variation per program. Ultimately, 61 prospective students participated in this study. They applied for six programs at four different universities in the Netherlands. The seventh program was left out, because only two students gave us permission to interview them. The participating prospective students varied with respect to age (18-24), gender (54% male), and type of previous education (75% high school, 20% university of applied science, 5% other/unassigned). The various types of matching for the different programs in our sample are displayed in Table 2.1.

Table 2.1 Type of matching per program.

	Type of program	Questionnaire	Online course	Interview	Matching day	Feedback or advice
Program 1 (U1)	non-STEM 1	CG	CG			CG
Program 2 (U2)	STEM	CG		CaR		CG
Program 3 (U2)	non-STEM 2	CG		CaR	VA	CG
Program 4 (U3)	STEM	CG	CG			CG
Program 5 (U3)	non-STEM 1	CG				
Program 6 (U4)	STEM	CG			CG	CG
Program 7 (U4)	non-STEM 2	CG		CR	CG	CG

Note: U1 = university 1, etc.; CG = compulsory activity for all students; CaR = compulsory activity for students at risk; VA = voluntary activity; CR = compulsory replacement of matching day for late applicants.

2.4.2 *Ethical Approval*

Permission for this study was obtained by the Ethical Review Board of Utrecht University's Faculty of Social and Behavioural Sciences (FETC17-098). Additionally, consent for recording the interview was requested and obtained from all participants. Participation was voluntary and confidential. Transcripts of the interviews were fully anonymized and cannot be linked to personal data or study progress.

2.4.3 *Procedures*

The researchers designed a topic list with sensitizing concepts derived from theory. The topic list consisted of three parts: standardized opening questions [e.g., *"what were your experiences with matching?"*], topics that were discussed according to applicability to the students' situation [e.g., *"what is your opinion about the advice the program gave you as a result of the matching?"*], and a standardized final question [i.e., *"what else do you want to address with regard to your program choice or matching in general?"*]. Follow-up questions were asked depending on the answers of the respondents and the type of matching they had completed. The phone interviews lasted approximately 20 minutes and the prospective students were informed of the expected duration beforehand. Following a grounded theory approach (Strauss & Corbin, 1994), data collection and data analysis occurred iteratively in the first stage of the data analysis.

2.4.4 *Analysis*

The data were analysed by the two interviewers using NVivo 11. Data analysis consisted of open, axial and selective coding techniques (Strauss & Corbin, 1994). First, nine interviews were conducted with prospective students from Programs 6 and 7. The two interviewers each employed open coding for all these interviews. Thereafter, they met to discuss three of these interviews. This meeting lasted until consensus was reached. During this meeting the decision was taken to slightly adjust the topic list. For example, during the introductory part of the interview the following question was added: "*did you also participate in other matching procedures?*" to avoid confusion about the program the interview concerned.

Then, the interviewers continued conducting the remaining interviews while recursively employing open coding. To ensure that consensus was reached, investigator triangulation (Merriam, 2009) was employed during the data analysis process. During this phase the interviewers met several times to compare and discuss the codes. These meetings lasted until consensus was reached. An example of something discussed during these meetings was the change from the concept "*expectancies*" to "*interests*". Initially, following the literal concept from EVT (Eccles & Wigfield, 2002), we assumed prospective students would have *expectancies* that would directly influence their program choice. However, when analysing the first set of interviews, emerging theoretical insights prompted us to change *expectancies* to *interests*, realizing that prospective students have expectations about many aspects of their future study (e.g., abilities). Therefore, the interviewers decided to update the topic list for the remainder of the interviews.

After the open coding phase, an experienced qualitative researcher was added to the research team to monitor the rest of the data analysis. Thereafter, the main researcher continued with the axial coding phase to condense and organize the data into meaningful categories. To increase trustworthiness in this stage, the main researcher made notes and memos. During the selective coding phase, a cross section was made between each element of matching and the theoretical concepts. By doing so we identified how each element played a role and how much impact each element had in the program choice. During this last phase, the category "*impact of the element*" was created by reallocating the codes about impact from the distinct theoretical concepts to a separate category. Throughout the data collection and analysis, the full research team met on a regular basis as an additional check on the research process.

2.5 Results

As stated in the theory, prospective students will choose a program if they experience positive feelings of fit with the program of their initial choice. It was hypothesized that prospective students can test this fit regarding abilities, interests and sense of belonging through the elements of matching. Table 2.2 shows the topics that prospective students mentioned per element of matching in positive (+) or negative (-) relation to these theoretical concepts. More specifically, prospective students indicated that their fit regarding *ability beliefs* and *interests* can be tested, to a greater or lesser extent, through all elements of matching. The theoretical concept *interests* was mentioned almost exclusively in positive terms in relation to the different elements of matching. Therefore, it appears that prospective students were best able to test their fit regarding *interests*. Furthermore, the only activity for which prospective students mentioned the possibility of testing *sense of belonging* was the matching day, thus being the only element that we considered to allow for testing *sense of belonging*. Therefore, the matching day is the only element of matching that was mentioned in relation to all aspects of person-environment fit. This implies that, from the elements of matching in this study, the matching day is perceived to allow for the most thorough check on person-environment fit.

Table 2.2 The possibility to test person-environment fit per element of matching.

	Questionnaire	n = 61	Online course	n = 15	Interview	n = 5	Matching day	n = 23	Advice		n = 52
Ability beliefs	Insight into time use	5	Confirmation of ability beliefs	10	Plays no role in assessing ability beliefs	2	(Part of) activities serve as ability check	20	Positive advice confirms ability beliefs	+	5
	Serves as ability check	3	Expects online course to be representative	6			Confirmation of ability beliefs	15	Negative advice results in doubt ability beliefs	+	3
	Does not play a role in checking ability beliefs	3	Expects online course to not be representative	6			Unsure about representativeness matching day	4			
Interests	No insights into content of the program	4	Insights into content of the program	8	Insights into content of the program	4	Insights into content of the program	17	Confirmation of interests	+	6
	Insights into content of the program	3	Insights into interest within program	5	Insights into interest within program	2	Confirmation of interests	17			
			Confirmation of interests	9							
Sense of belonging							Feeling at home	11			
							Connection with people	13			
							Good atmosphere	8			
							Cannot judge feeling at home	4			

2.5.1 *Statements Confirming the Possibility to Test Person-Environment Fit*

Confirming statements about the possibility to test fit with a program are the result of matching elements allowing prospective students to test whether they feel able, are interested in the materials/topics of the program and experience a sense of belonging. First, a confirming statement often made regarding abilities is that the matching serves as an ability check. This was mainly the case for the matching day and less so for the questionnaire. In relation to the matching day, it was mentioned that quizzes, homework, lectures or tests made it possible for prospective students to test whether they would be able to keep up with the level of the program. Related statements came from prospective students who did an online course; some of them mentioned that they expected the level of complexity of the online course to be representative for that of the actual program. Many prospective students experienced the content studied during the online course or matching day as a confirmation of their ability beliefs. This confirmation was also found by some prospective students in the advice they received based on the matching activities.

> Yes, that [positive advice] was a relief, that it is a confirmation of yes, you can handle the program.
> *Program 1, student 10*

Confirming statements for testing person-environment fit regarding interests were made in relation to all elements of matching. The statement most often made was the fact that the matching element gave insight into the content of the program, either through filling out the questionnaire, talking to a staff member, doing the online course or by participating in the matching day. Prospective students participating in interviews and online courses sometimes mentioned that these activities gave them insights into specific interests within the program content. Many prospective students, especially those who participated in an online course or a matching day, said that the matching gave them a confirmation that they found the program interesting.

> I think I am going to like the program very much. I was already enthusiastic because of the assignments [in the online course], that represented the content of the program.
> *Program 4, student 1*

Lastly, almost all prospective students that participated in a matching day mentioned at least one of the following aspects that made them feel a sense of belonging during the matching: feeling at home on campus; having a good connection with staff and other prospective students; feeling a good atmosphere on campus.

> Because the day before I had done the matching of [another Bachelor program]. But I also like that very much, but when I came to [Bachelor program's Matching day] I really had a feeling like, that I came home. It really suits me.
> *Program 7, student 3*

One prospective student missed this connection with fellow prospective students, staff, current students, and discussed topics, and therefore concluded that the program was not the right fit. This student's decision to opt for another program was mostly based on lacking feelings of social fit.

Prospective students whose statements confirmed the possibility to test fit on one of the theoretical concepts, often indicated the possibility to also test fit on another theoretical concept. However, some prospective students were actively searching for confirmation regarding just one aspect. For example, some students were mainly talking about the level of the matching procedure, because they were actively looking for a confirmation of their ability beliefs.

2.5.2 Statements Disconfirming the Possibility to Test Person-Environment Fit

Although much less than the positive statements, also negative statements regarding the possibility to test person-environment fit were made. These statements represent a failure to assess abilities, interests or sense of belonging through the matching elements. Failure to test ability beliefs was mentioned several times, either because the element (i.e., interview) did not play a role in checking abilities (not the right means) or because the element (i.e., online course or matching day) seemed not representative for the program content (not the right materials). Some prospective students of Program 1 were quite outspoken in the fact that they struggled with the representativeness of the material in the online course.

> Well, yeah, I don't know if…. well, in the matching procedure I had… I didn't really feel like the questions were at the level, like, what you have to be able to do at the university and actually I'm a little curious, so to say, what the level would be, yes,… that is also quite difficult. Yes, well, actually more whether I can handle it, indeed, because I had no idea, say, at what level you can make a choice.
> *Program 1, student 7*

Furthermore, some students criticized the questionnaire for not being able to give insights in the content of the program (not the right means) and others said that one matching day is too short to get a sense of belonging. Thus, these students did not find the matching sufficient to test their person-environment fit thoroughly.

2.5.3 *Contradictory Statements Regarding the Possibility to Test Person-Environment Fit*

Throughout the data analysis, it became apparent that for some matching elements different subgroups have varying views on the possibility to test their person-environment fit. First, the online module was experienced differently by prospective students from the two programs that offer this element of matching. The prospective students who applied to Program 4 all found confirmation of their ability beliefs through the online course. Those who applied to Program 1 generally found the online course easy or not difficult. More than half of them got the feeling that the level of the course was not representative of the program. Some of these students therefore said that this made it difficult to estimate the level of the program and that they gained no insight into whether they would be able to cope with the level of the program. It thus seems to be necessary for prospective students to experience the matching as representative to find it a useful means to test their fit. The second difference is between STEM students and non-STEM students in their experience of the matching day. Most prospective students applying to a non-STEM program expressed that they felt comfortable with the general atmosphere on campus and specifically with the program. They experienced both staff and current students as open and friendly with a welcoming attitude towards prospective students. Although prospective STEM students also expressed feeling a positive atmosphere at the university, many of them mentioned not feeling comfortable in the beginning of the day but soon realized that their fellow prospective students had similar interests. On the other hand, some of them explained that they had difficulty assessing within a day whether they felt at home socially. This indicates that sense of belonging might play a different role for non-STEM students than for STEM students when testing their person-environment fit.

2.5.4 *Impact of Matching on Program Choice*

Apart from looking at the theoretical concepts of fit, we also asked our respondents about the impact of the matching elements on their program choice. Results are displayed in Table 2.3. There are clear differences in the impact of matching elements on prospective students' program choice. The interviews had the biggest impact on their program choice. Most prospective students expressed that they experienced the personal contact with a staff member as valuable. They all said that the interview gave them confirmation of their program choice. Two prospective students pointed out that, though they were already quite certain about their choice, the opinion of the staff member added to their feelings of certainty. Thus, "qualitative" feedback from staff members seems important in prospective students' choice.

> I felt like the study would suit me, but with the matching activity that you get to hear from an instructor who teaches there that they think the study suits you, that gives extra confirmation.
> *Program 7, student 1*

The student-staff interviews were only conducted with a small group of students who were deemed at risk of dropping out. A larger number of prospective students participated in an online course or matching day. As shown in Table 2.3, about half of the prospective students indicated that the matching day served as a confirmation of their program choice. For the matching day, somewhat less than half said that it had no impact on their program choice, which is, in part, caused by the fact that some of them were already certain about their choice. The impact of the online course on program choice differed for the prospective students who were very positive about the online course and experienced it as representative (Program 4) and those who experienced it as (too) easy and not representative (Program 1). This appears to reflect the extent to which they felt that they could test their *ability beliefs*. All prospective students who applied to Program 4 said that the online course served as a confirmation of their program choice. However, few prospective students of Program 1 said that the online course gave them choice certainty and about half of them indicated that the online course did not contribute much to their program choice.

The last two elements in the matching procedures are the questionnaire and advice. Although every prospective student starts the matching with a questionnaire, not many stated that it impacted their program choice. In fact, just over half of the prospective students said that it had no value for their program choice. Most of these students had not participated in any activities other than the questionnaire. Additionally, some students, such as the student quoted below, did not remember they had filled out the questionnaire. In contrast to those whose matching consisted of only a questionnaire, of the prospective students who also participated in other activities as part of their matching less than half indicated that the questionnaire had no value.

> I suspect I did. I think a while ago, but I can look for it, but as far as I'm concerned, we'll start the interview. ...but yes, I'm really thinking, what was that matching questionnaire about? I think I did it two months ago or so. And now I can't find it unfortunately.
> *Program 5, student 4*

Table 2.3 The impact of the matching elements in testing person-environment fit.

	Questionnaire	$n = 61$	Online course	$n = 15$	Interview	$n = 5$	Matching day	$n = 23$	Advice	$n = 52$
Impact of the element in choosing a program	No added value in program choice	31	Confirms program choice	7	Confirms program choice	5	Confirms program choice	11	No added value in program choice	14
	Sparks thinking about reasons for choosing program	17	No added value in program choice	6	Personal contact is valuable	3	No added value in program choice	8	Positive advice has no influence on program choice	14
	Adds to last bit of choice certainty	5							Positive advice confirms program choice	8
	Helps structuring thoughts	5							Ignores negative advice	4
									Would (maybe) ignore negative advice	6
									Would (maybe) doubt program choice in case of negative advice	5

2

That prospective students who only filled out the questionnaire felt less capable of testing their fit also became clear from the finding that all statements about ability beliefs and interests that disconfirmed the possibility of checking fit were made by this group. They also indicated less often that the questionnaire sparked thinking about reasons for choosing the program. Even though most of this group felt less capable of testing their fit, a few students from this group indicated that the questionnaire gave them the choice certainty they needed and that it helped them structure their thoughts.

Notably, the advice was perceived as even less relevant in program choice than the questionnaire. Hardly any students indicated that the advice impacted their program choice or could potentially do so. However, also for this matching element perceptions of impact differed between subgroups. We found that more than half of the prospective students who applied for Program 1 said that the advice had no value for their program choice. This is substantially higher than at the other programs, where only a few prospective students said that the advice had no impact on their program choice. Moreover, at Program 1, some students said that they ignored a negative advice. The most likely explanation for taking the advice less seriously is that they perceived the online course as non-representative of the actual program. Negative advice can potentially make prospective students doubt their choice. Some prospective students stated that they would have questioned their program choice if they had received negative advice.

> I mean, if it was a negative result like, we do not think that this program suits you, then I would certainly have started to doubt whether I had made the right choice.
> *Program 1, student 3*

Some prospective students at several universities affirmed that positive advice had no influence on their program choice or that they would ignore negative advice. They indicated that they were certain about their choice and that, therefore, the advice had no value.

Overall, it seems that the better an element of matching allows for testing person-environment fit regarding abilities, interests and sense of belonging, the more impact this element of matching has on the final program choice. Given the small and specific group of prospective students that did an interview, it is difficult to say something about the impact of the interview on program choice. However, the online course and matching day both allow for a substantial assessment of person-environment fit and seem to have impact on program choice. The questionnaire and advice, on the other hand, seem less suitable for testing fit and are also not perceived as having much impact on students' program choice.

2.5.5 Choice Certainty

Choice certainty emerged from the data as a core category that influences how prospective students experience the possibility to test fit and the impact matching has on their program choice. It appears that the more certain they are about their initial choice, the less impact the matching has on their final choice and, thus, the less likely a change in choice certainty will take place. However, some prospective students who claimed to be 100% certain about their initial choice nevertheless said that the matching made them more certain about their final program choice. On the other hand, some prospective students indicated that the online course and the matching day did not impact their program choice. All prospective students who reported that the online course had no value, and about half who reported that the matching day had no value, stated that they were already certain about their choice. Hence, the level of choice certainty during the initial program choice influences how much impact the matching can have on final program choice and the certainty with which students make this choice. Given that we found that person-environment fit can be tested, at least partially, through all elements of matching, we propose an integrated framework for program choice as shown in Figure 2.2.

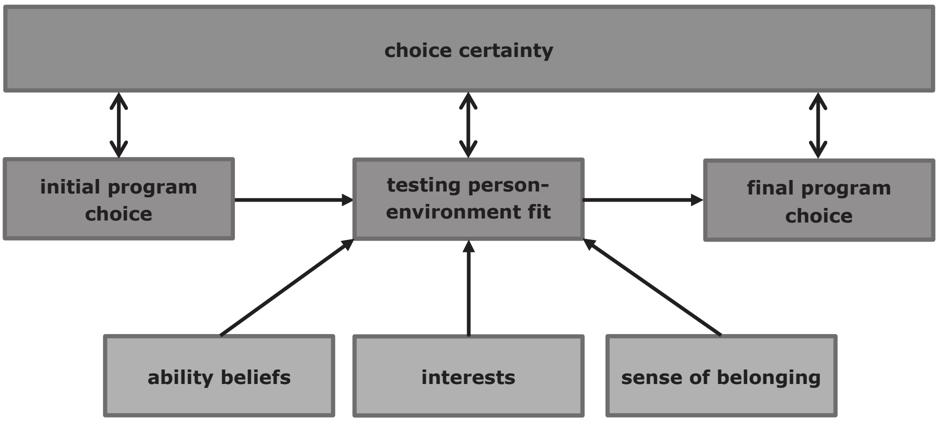

Figure 2.2 Assessment of person-environment fit in choosing a program as emerged from the data.

2.6 Conclusion and Discussion

Prior work on the transition to higher education has studied program choice from specific theoretical angles (e.g., SDT, Deci & Ryan, 1985; EVT, Eccles & Wigfield, 2002). We proposed a more comprehensive theoretical framework on person-environment fit combining several theoretical perspectives to look at the choice process in the transition to higher education. In addition to testing person-environment fit by testing fit on interest, ability beliefs and sense of belonging, the framework with choice certainty as an overarching core category as represented in Figure 2.2, adds to the theoretical understanding of the transition from secondary to higher education. Prospective students who are participating in matching procedures are looking for insights into the program as well as for confirmation of their program choice.

In studying matching procedures in the context of this framework, we found that the extent to which matching procedures are the right means for achieving these two goals differs both between elements of matching and between students within one element of matching. Prospective students perceived all elements of matching as useful, to a greater or lesser extent, to assess person-environment fit for the program. The questionnaire and advice were not mentioned often in general and when mentioned, they seemed to have little impact on program choice.

The other three elements of matching that we investigated, interview, online course and matching day, are perceived as more useful in testing fit on interest and ability beliefs. The only element in our study that allows for testing fit regarding the three theoretical concepts is the matching day. Research shows that the possibility to test fit on sense of belonging seems especially important for successful adjustment to university (Brinkworth et al., 2009; Mattanah et al., 2010). Having been on campus before starting the program makes this adjustment easier. However, it could be argued that sense of belonging can also be tested before making an initial program choice. Several prospective students indeed mentioned that they did not feel the need for a matching day, because they already visited the campus (e.g., during an open day). In such instances, testing sense of belonging might not have to be part of the matching process. In that case, an online course may be just as useful to test person-environment fit and is, in the long run, less labour intensive for university staff. Nonetheless, staff should be aware of whether every prospective student had the opportunity to visit campus and whether there are other advantages of being physically present on a matching day, such as personal contact with staff and current students.

Furthermore, we showed that certain elements are experienced differently by various prospective student groups. The online module was experienced as a better means to test fit by prospective students from program 4 than by prospective students from program 1. This appears to be the result of prospective students of program 1 experiencing the online module as not representative for the program. Hence, the (experience of) representativeness is an important aspect to consider. Furthermore, based on the experiences of prospective students with the matching day, it seems that sense of belonging plays a different role in testing person-environment fit for STEM students than for non-STEM students. This implies that particular student groups may benefit from varying types of matching procedures. It is important, however, to be aware of the fact that other factors than those identified here could have played a role in causing these differences.

Altogether, it seems to be useful for prospective students to experience the program of their initial choice before making a final choice. In experiencing the program, prospective students can test their person-environment fit. Dutch matching procedures seem to allow for such a fit test. However, there are differences between these procedures in the extent to which they allow for testing fit and the extent to which they are experienced as useful in making a final program choice. Based on the conducted interviews it can be cautiously stated that the more elements of person-environment fit can be tested in a matching procedure, the more it impacts the prospective students' final program choice. However, a more extensive matching procedure is not necessarily a better means to test fit. It seems vital for the usefulness of matching in program choice that prospective students experience the content as representative. On top of that, not all matching procedures might be equally useful for all student groups (i.e., STEM vs. non-STEM and other student characteristics). Therefore, program staff should be aware of the type of student entering their program.

There are several limitations to this study that need to be mentioned. First, this study focused on perceptions of prospective students in assessing the usefulness of different elements of matching for testing fit. Factual data on, for example, academic success was not included, and therefore, it is not possible to determine whether the elements of matching that were perceived as useful are the ones that are effective in increasing academic success. In the future, our results can be compared to quantitative academic achievement data. Second, we only conducted 5 interviews with students who participated in an interview as part of their matching, because few prospective students were requested to participate in such an interview. Although generalization is not an aim of this study, the low number of interviews should be considered when reflecting on the usefulness of an interview for testing fit. Lastly, we did not consider whether the elements we studied were compulsory. There might be differences in experiences and impact between those who participated in matching voluntarily and those who participated in compulsory activities.

All in all, our research shows that some elements of matching help prospective students to test their person-environment fit for the program of their initial choice. Using a matching procedure where prospective students can test different aspects of person-environment fit may benefit both prospective students and higher education institutions in helping more students to make a validated program choice and ultimately lead to less drop out in higher education. It is known that congruence between student and program is associated with academic success (Feldman et al., 1999). Higher education institutions in other European countries could consider implementing similar procedures for open-admission programs to foster prospective students' program choice. Future research should be aimed at broadening the empirical basis and investigate whether the matching elements are related to academic success.

Chapter 3

Determining Fit: The Role of Matching Procedures in Prospective Higher Education Students' Enrolment Behavior

Soppe, K. F. B., Klugkist, I.G., Wubbels, T., & Wijngaards-de Meij, L. D. N. V. (2022). Determining Fit: The Role of Matching Procedures in Prospective Higher Education Students' Enrolment Behavior.

ABSTRACT

Students experiencing a misfit between themselves and their university program can result in various negative consequences. Hence, improving the process of program choice might foster student-program fit. In the Netherlands, the implementation of mandatory enrolment procedures, in which prospective students do a final check on their initial program choice (so-called matching procedures), were introduced to improve the student-program fit. We argue that prospective students who lack feelings of fit during these matching procedures are less likely to finalize their enrolment. Using data of thirteen programs at four Dutch universities, the association between a wide array of matching procedures and finalizing enrolment was examined. Enrolment rates are lower in programs with more intensive matching procedures. Moreover, enrolment rates are higher in matching cohorts than in pre-matching cohorts, indicating the potential value of pre-enrolment fit-checks. In conclusion, this study gives indications that it can be worthwhile to invest in guiding prospective students in their program choice by obliging them to test fit with the program through intensive matching procedures.

Author contributions: KS, LWM, TW, & IK designed the study. KS & LWM recruited partner universities. KS organized the data collection, conducted the data analysis, and wrote the paper. TW, IK, & LWM provided extensive feedback on the manuscript.

3.1 Introduction

Students in higher education who do not experience fit with the program they choose are more likely to drop out from university than students who do experience student-program fit (Feldman et al., 1999). Feelings of misfit become clear when students, once they started studying, realize that their expectations do not fit with the reality of the program (e.g., Warps et al., 2017, p.11), which could result in dropout once more realistic beliefs set in (Watson et al., 2004). When students have the chance to form realistic beliefs regarding their program fit prior to enrolment disappointment in the program can be prevented.

In fact, several European studies have shown the need to improve the process of program choice of prospective students to advance the match between student and program (see for example: Austria (Unger et al., 2009), Flanders (Goovaerts, 2012), Germany (Heublein et al., 2010), the Netherlands (Meeuwisse et al., 2010), and Switzerland (Wolter et al., 2013). When students are being guided in their program choice, universities are supposed to end up with a better fitting student population.

It is thus important that prospective undergraduate students test their fit with the program before actually starting a program in order to reduce non-fitting program choices. This study explores the association between fit-checks and enrolment rates, using data of different types of so-called matching procedures in the Netherlands. These matching procedures aim to offer a fit-check prior to enrolment. If these procedures function as envisaged, students who experience a lack of fit, will drop out of the enrolment process rather than the program itself. Hence, finding an association between matching procedures and enrolment rates, provides an indication that fit-checks prior to enrolment can advance the match between student and program.

3.2 Person-Environment Fit

In this study, data on the enrolment behavior of prospective undergraduate students are analyzed within a theoretical framework on person-environment fit. Person-environment fit is the compatibility between individual and environmental characteristics (Kristof-Brown & Guay, 2011). Fit research across a variety of domains shows that an individual's performance improves if there is alignment between a person and their environment (Ward & Brennan, 2020). Drawing on insights from Expectancy Value Theory (e.g., Eccles & Wigfield, 2002; Wigfield & Eccles, 2000) and Self-Determination Theory (Deci & Ryan, 2008), we argue that to be able to test fit, prospective students should be able to: 1) verify the accuracy of their ability believes (i.e., does the extent to which they believe they are capable of performing a task match with reality (Bandura, 1977; Deci & Ryan, 2008)); 2) identify to what extent they value the content (i.e., do their interests match with the content of the program (Wigfield, 1994; Wigfield & Eccles, 1992)); and 3) experience the social aspects of the new environment (Tinto, 1987), because the type of students and staff involved is also an inherent part of identifying with and belonging to a higher education program (Hoffman et al., 2002; Swenson et al., 2008). Hence, we define testing student-program fit as students asking themselves whether they match with the program of their choice regarding ability beliefs, interests and sense of belonging.

A fitting program choice is associated with positive outcomes for the student. For example, experiencing a sense of belonging positively influences students' engagement in the academic process (McFarlane, 2018; Trowler, 2010) and is associated with a lower probability of dropout (Kirk, 2018). Previous work highlights that experiencing feelings of misfit in general (Feldman et al., 1999; Ulriksen et al., 2010; Warps et al., 2017) or more specifically in sense of belonging (Naylor et al., 2018; Tinto, 1987) are among the most important predictors of dropout.

3.3 Matching Procedures in the Netherlands

The enrolment procedure for prospective undergraduate students wishing to enroll in an open-admission Dutch university program consists of several steps. The main steps are visualized in Figure 3.1. After filing an initial admission request, prospective students are invited to participate in a matching procedure in order to check whether they made the right choice (Association of Universities [VSNU] n.d.). Thereafter, students can either finalize their enrolment for the program of their choice or opt for another program or institution or decide not to enroll at all. Prospective students who did not request admission before the deadline, or those who did not take part in the matching procedures can be denied admission by the program staff. All higher education institutions in the Netherlands are obliged to offer prospective students applying for open-admission programs the possibility of taking part in a matching procedure (Wet Kwaliteit in Verscheidenheid Hoger Onderwijs [Quality in Diversity Law] 2013). These procedures differ from selection in that selection is binding, while matching procedures are not: it remains the student's decision whether to enroll or not.

In this study we explore whether matching procedures, that were implemented to improve students' program choice are related with student enrolment behavior, by studying whether and how different types of matching procedures are associated with enrolment rates. To be useful as a fit-check, matching procedures should allow for testing fit on at least one of the three components of student-program fit discussed above. Earlier research shows that this is the case to some extent for all types of matching procedures offered at Dutch universities (Soppe et al., 2019). Moreover, Niessen and colleagues (2016) have shown that these matching procedures can be useful for preventing a wrong choice. Using data from one Dutch university, they showed that the lower students' scores on a curriculum sampling test at the end of the matching procedure, the less likely they were to finalize their enrolment. The implementation of matching procedures is an attempt to let non-fitting students drop out during the enrolment procedures (i.e., realize there is a lack of fit before they start studying).

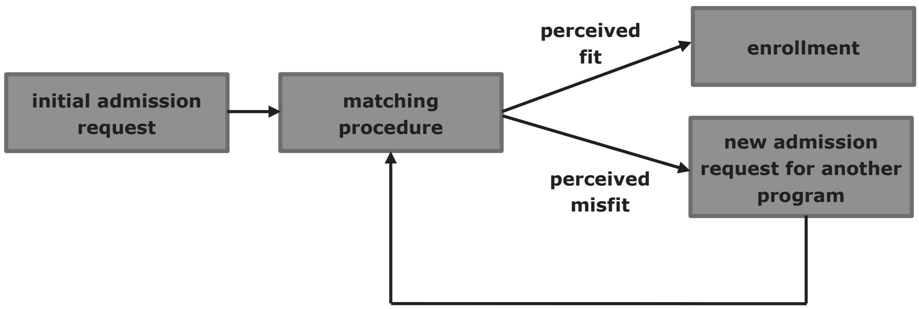

Figure 3.1 Enrolment procedure for prospective students at Dutch universities.

Matching procedures can take different forms, because universities were entitled to develop these themselves, according to their own needs and insights. Across and within universities there is therefore a wide variety of procedures, but there are several elements that are offered at almost every university. Most matching procedures start with an online questionnaire. Questionnaires are not the same across universities, but generally contain similar sections such as (a selection of): high school grades; attendance at orientation activities to learn about different programs; motivation or reasons for choosing the program; expectations of the program; expected future jobs; ability beliefs; general time use. Most matching procedures end with generic feedback or nonbinding advice on the perceived fit between a prospective student and the program. In between these two elements of the procedure, three main activities can be identified. Some programs offer personal interviews to prospective students who are deemed at risk of dropout based on their answers on the questionnaire (type 1). Other programs offer short online courses (type 2). And finally, there are programs that offer trial studying on campus (type 3).

Prospective students experience that certain types of matching allow for more thorough testing of student-program fit than other types (Soppe et al., 2019). Trial studying on campus (type 3) is perceived as the most thorough check of student-program fit among the three types of matching procedures that will be considered in this study. A type of matching that was perceived almost equally helpful in testing fit was the online course (type 2), which generally consists of video lectures and tests. However, in type 2 matching procedures testing sense of belonging is not easily possible, given the individual, online set-up. Type 1 matching procedures, which consist of interviews with students deemed at risk of dropout, are considered least helpful by prospective students in testing fit. Given these results, it is important to investigate whether the experiences reported by students on testing student-program fit are also reflected in the enrolment rates (i.e., the ratio between the number of students that files an initial admission request and the number of students that finalizes their enrolment) across programs with different kinds of matching procedures as well as differences in enrolment rates between cohorts before and after the implementation of the matching procedures.

In this study we look at enrolment rates at the program level. The lower the enrolment rate, the more prospective students have chosen not to finalize their enrolment after participating in a matching procedure. This implies that lower enrolment rates are positive for the program, if indeed the matching procedure helped students changing their mind when experiencing feelings of misfit during the matching procedure. Only prospective students who do experience fit are expected to enroll.

By comparing enrolment rates between different types of matching procedures at four Dutch universities, we identify whether enrolment behavior differs across these types of matching. Our research question is: *How do enrolment rates vary between university programs with different types of matching procedures? How do these rates differ before and after implementing matching procedures?*

This overarching research question will be answered in four steps. First, to establish an overview of enrolment rates across university programs with different types of matching procedures, the data of matching cohorts will be used. Second, in order to identify and isolate the influence of possible university effects, at one university with different types of matching procedures we examine the enrolment rates per program. Third, we determine whether enrolment rates differ across academic disciplines with different types of matching procedures, regardless of the university at which they are offered. Fourth, we establish whether these potential differences in enrolment can be attributed to the implementation of the matching procedures, by comparing cohorts with matching procedures to cohorts before the implementation of matching.

3.4 Methods

3.4.1 Sample

The sample includes information on academic records (N=20,104) of nine cohorts (2009 – 2017) of students who applied at four different academic disciplines (A, B, C and D) at four Dutch universities (thirteen study programs in total). Discipline C is a discipline in the domain of Science, Technology, Engineering and Mathematics (STEM). Disciplines A, B and D are classified as social and/or human sciences. All programs are matching programs (i.e., none of the programs applies selection procedures or capacity constraints). The records of the students in our sample contain information about enrolment and the type of matching procedure. An overview of the available data per cohort and types of matching procedures is displayed in Table 3.1.

Table 3.1 Available data per cohort; pre-matching cohorts in grey.

Cohort	**University**			
	U1	U2	U3	U4
Type of matching	Mixed[1]	Type 1	Type 2	Type 3
2009				ABCD
2010				ABCD
2011				ABCD
2012	ABCD			ABCD
2013	ABCD			ABCD
2014	ABCD	ABCD		ABCD
2015	ACD[2]	ABCD		ABCD
2016	ABCD	ABCD		ABCD
2017	ABCD		C[3]	ABCD

Prospective students (55% male) were on average 19 years old (*M*=19.0; *SD*=2.0) during their application. Those who applied to a program in 2014-2017 belong to a matching cohort, as do prospective students who applied to University 4 in 2013. The remaining students applied to a program before the implementation of match-

[1]Program A and D applied type 3 matching, program B applied type 2 matching and program C applied type 1 matching.

[2]Data for program B of university 1 is missing since selection procedures were applied for this cohort.

[3]Obtained data for University 3 consists of two programs across five cohorts. Only the registration of admission data for program C in cohort 2017 was suitable for calculating enrolment rates. Because it is a large program that represents the online matching in our sample (together with program B of U1), it was decided to keep this information in the dataset, despite it being the only usable data of University 3.

ing (hereafter pre-matching cohorts). Participation in the matching procedures was a prerequisite for finalizing enrolment in all programs. For students participating in type 1 matching procedures this entails filling out the compulsory questionnaire, potentially followed up by an interview. Students (N=992) with uncommon applications (e.g., those with an exemption for the matching procedures), often follow non-standard enrolment procedures and were not included in the analyses.

3.4.2 Procedure

Admission and matching information was drawn from the registration systems by data managers of each university. All information was anonymized before being transferred to the authors. Permission for this study was obtained by the Ethical Review Board of Utrecht University's Faculty of Social and Behavioural Sciences (FETC17-098).

3.4.3 Analyses

The nature of this study is exploratory, and analyses are therefore mainly of a descriptive nature. Assumed differences in the enrolment rates across the three types of matching were tested for significance using chi-square association tests.

This study combines data from different universities which allows for a comparison of the three main types of matching procedures. Since these data are acquired from the different university registration systems, they are different in nature. As a result, analyses on the individual level, commonly used in this kind of research, are not possible with our data. However, since we are interested in the association between matching procedures and enrolment rates, individual-level data are not required to answer our research question.

The analyses were carried out in four steps. First, differences in enrolment across programs with different types of matching were identified (data of matching cohorts, N= 11,404). Second, to explore the influence of institutional variation, the comparative analyses were repeated using data of one university with different matching procedures (University 1: N = 3,517, see Table 3.1). Third, the data was presented per academic discipline instead of per university to identify differences in the enrolment rates of programs within the same discipline across different universities (N = 11,404). In the fourth step, differences between pre-matching cohorts and matching cohorts were investigated. Due to differences in registration of matching data across institutions, this final step is conducted for University 1 (N = 6,201) and University 4 (N = 10,533).

In each of the four steps enrolment rates were visualized in a figure and significance of differences was tested using chi-square tests, including post-hoc tests. Results of significance and effect sizes for the post-hoc tests for the first three steps are presented in Table 3.2. Results of significance and effect sizes for the post-hoc tests for the comparison between matching and pre-matching cohorts (step 4) are presented in Table 3.3.

3.5 Results

In the first step, enrolment rates across types of matching were compared for all matching cohorts and disciplines combined. Figure 3.2 shows that for the total sample, enrolment rates were highest for programs with type 1 matching (interviews for students at risk of dropout) and lowest for programs with type 3 matching procedures (matching day on campus). The association between enrolment and type of matching is significant, χ^2 (2, N = 11,404) = 128.56, p <.001, but weak, V = .11 (Cohen, 1988). Two post-hoc tests also showed significant results. Enrolment rates were higher in programs with type 1 matching procedures than in programs with type 2 matching procedures and enrolment rates were higher in programs with type 2 matching procedures than in programs with type 3 matching procedures.

Second, Figure 3.2 also shows enrolment rates per university. For University 1, enrolment rates followed the general pattern. University 2, 3 and 4 each applied one type of matching procedure. Enrolment rates were highest in University 2 (type 1 matching procedures) and lowest in University 4 (type 3 matching procedures). The different matching procedures within University 1 made it possible to look at variations in enrolment rates between types of matching within one university, therefore cancelling out the influence of specific university characteristics. Results of testing the association between type of matching and enrolment rates on the data of University 1 alone showed a significant, χ^2 (2, N = 3,517) = 15.90, p <.001, but weak association between type of matching and enrolment rates (V = .07). Within University 1, enrolment rates were higher in programs with type 1 matching than in programs with type 2 matching. Differences in enrolment rates between type 2 and type 3 matching procedures, however, were not significant.

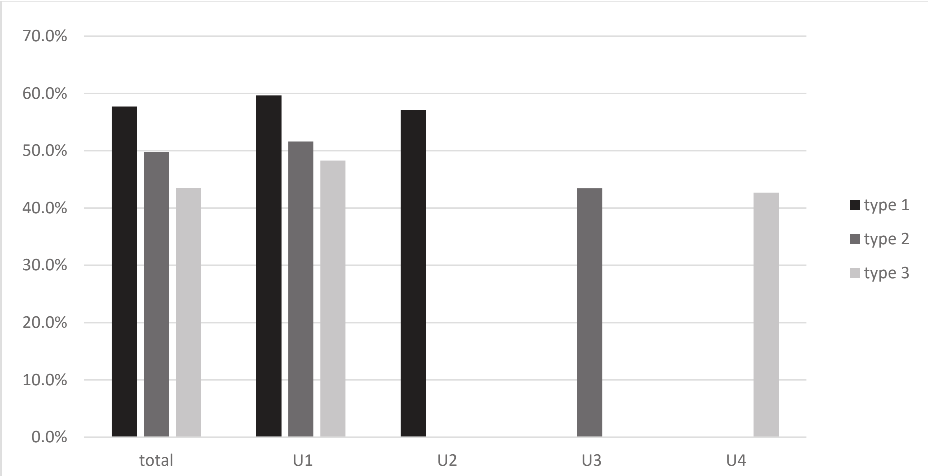

Figure 3.2 Enrolment rates by type of matching in total and per university; all matching cohorts combined.

Third, Figure 3.3 shows the association between type of matching and enrolment rates per academic discipline. The association between type of matching and enrolment rates was significant for all programs of discipline A combined, χ^2 (2, N = 1,901) = 21.19, p <.001, V = .11. Enrolment rates for discipline A were higher in programs with type 2 matching procedures than in programs with type 3 procedures, while there was no difference in enrolment rates between type 1 and type 2 matching procedures. For all programs of discipline B combined, the association was also significant, χ^2 (2, N = 5,575) = 110.03, p <.001, V = .14. Enrolment rates for discipline B were higher in programs with type 1 procedures than in programs with type 2 procedures and higher with type 2 procedures than in programs with type 3 procedures. For all programs of discipline C (STEM) combined, there was also a significant association between type of matching and enrolment, χ^2 (2, N = 2,287) = 36.71, p <.001, V = .13. The pattern of enrolment rates for this program deviated from the general pattern. Enrolment rates for discipline C in programs with type 2 procedures were lower than enrolment rates in programs with both type 1 and type 3 procedures. Finally, none of the programs of discipline D in our sample employed type 2 matching procedures. Enrolment rates for discipline D in programs with type 1 procedures were, however, higher than enrolment rates in programs with type 3 procedures, χ^2 (1, N = 1,641) = 14.82, p <.001, V = .10.

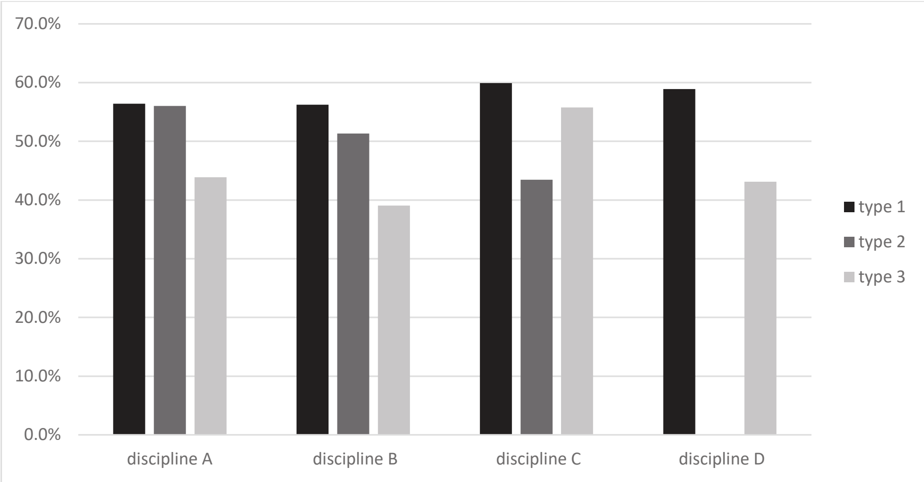

Figure 3.3 Enrolment rates by type of matching per discipline; all matching cohorts combined.

To sum up, there were generally small but significant differences in enrolment rates across programs with different types of matching procedures. Overall, enrolment rates were lowest in programs that employ matching days on campus and highest in programs with interviews for students at risk of dropout. Programs in the STEM domain appeared to be an exception to this overall pattern. Here, enrolment rates were found to be lowest in the program with an online course, rather than the program with a matching day.

Table 3.2 Post-hoc tests showing the association between enrolment and type of matching procedures for the total sample and different sub samples.

	Type 1 vs. Type 2 procedure	Type 2 vs. Type 3 procedure
Total sample	χ^2(1, N = 4,582) = 28.04, p <.001, V = .08	χ^2(1, N = 9,504) = 30.62, p <.001, V = .06
University 1	χ^2(1, N = 2,546) = 9.65, p =.002, V = .06	χ^2(1, N = 3,066) = 2.89, p =.089, V = .03
Discipline A	χ^2(1, N = 444) = .01, p =.945, V = .00	χ^2(1, N = 1,573) = 6.44, p =.011, V = .06
Discipline B	χ^2(1, N = 2,742) = 5.28, p =.022, V = .04	χ^2(1, N = 4,812) = 71.49, p <.001, V = .12
Discipline C	χ^2(1, N = 1,233) = 33.41, p <.001, V = .17	χ^2(1, N = 1,641) = 23.01, p <.001, V = .12

In the fourth step, changes in patterns of enrolment rates over time were investigated. Pre-matching data was available for University 1 and University 4. Figure 3.4 shows the enrolment rates over time per discipline for University 1. No difference was observed in the average enrolment rate before and after the implementation of matching for disciplines that offered matching activities only for students that were deemed at risk of dropout (discipline B and C). On the other hand, programs of disciplines which offered intensive matching procedures for all students allowing to test (almost) all aspects of student-program fit (disciplines A and D), showed considerable differences in the average percentage of students that finalized their enrolment before and after matching was implemented. Effect sizes for these associations over time in disciplines A and D are medium-small. Significance and effect sizes are shown in Table 3.3 to aid comparability between universities.

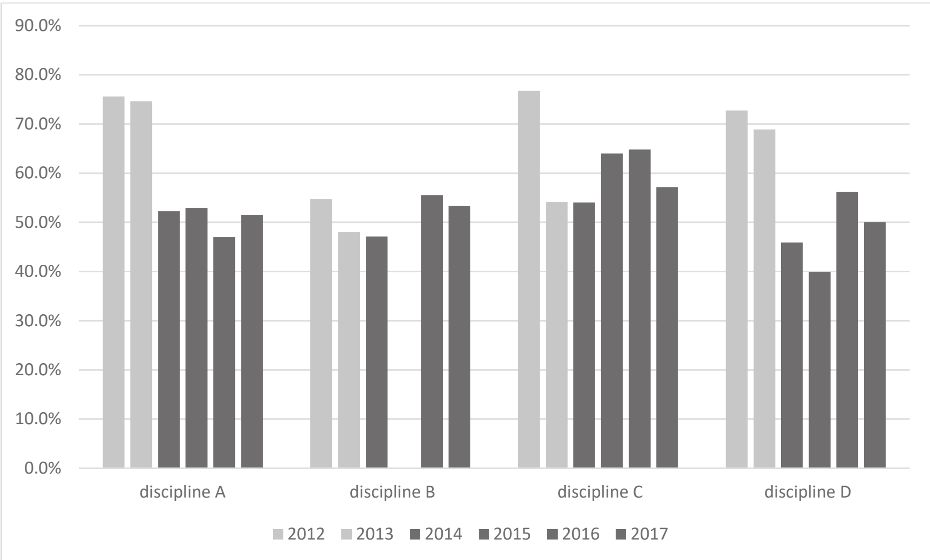

Figure 3.4 Enrolment rates per program of University 1[4] before (light grey) and after (black) the implementation of matching.[5]

[4]All programs of University 4 applied type 3 matching procedures. For this university that means that prospective students start by filling out an online questionnaire, subsequently prepare for a matching day through means of homework, attend the matching day and conclude that day with a test. After the matching day they receive an email containing feedback, but no concrete advice.

[5]Data for Program B is missing for the cohort 2009.

The results of the enrolment rates over time for University 4, which employed intensive matching procedures for all disciplines (type 3), are displayed in Figure 3.5. There was a clear drop in enrolment rates before and after the implementation of the matching procedures for discipline A, discipline C, and discipline D. Effect sizes for these associations over time are medium-small to medium. Enrolment rates for discipline B were slightly higher after the implementation of the matching procedures, but the effect is small.

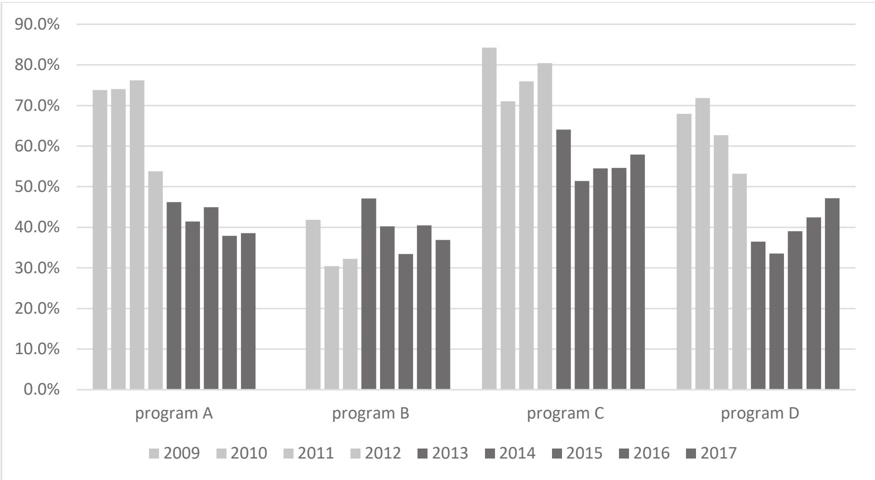

Figure 3.5 Enrolment rates per program of University 4[6] before (light grey) and after (black) the implementation of matching.[7]

[6]All programs at University 1 start the matching procedures with an online questionnaire and conclude the procedures with an advice (red, orange, green – traffic light analogy) on the expected fit of a prospective student. The matching procedures per program for the cohorts in our sample are as follows: Prospective students who applied for Program A had the possibility to participate in one of the two matching days, or alternatively follow the online matching procedure. Program B offered a short online course for students deemed at risk of drop-out based on their answers to the questionnaire. Program C offered students that were deemed at risk of drop-out a personal interview. The matching procedure for Program D consisted of a matching day, including group interviews.

[7]Data for program B of University 1 is missing since selection procedures were applied for this cohort.

The available data for University 4 shows that the number of prospective students who filed more than 1 initial admission request within the same university roughly doubled for all programs since the implementation of the matching procedures. Since an increase in the number of prospective students filing multiple initial admission requests would automatically result in lower enrolment rates (because the vast majority of students only starts with one program), the analyses for University 4 were rerun on a subsample with only the students that filed 1 initial admission request (see Table 3.3). The results are the same for University 4 as a whole and the subsample, indicating that an increase in the number of prospective students filing multiple initial admission requests does not explain the changes in enrolment over time.

Table 3.3 Chi-square tests of the association between belonging to a matching cohort and enrolment for the different disciplines in University 1 and University 4.

	Total sample U1 with all admission requests	Total sample U4 with all admission requests	Sub sample U4 with only applicants with one admission request
A	$\chi^2(1, N = 860) = 51.84$, p <.001, $V = .25$	$\chi^2(1, N = 2380) = 173.37$, p <.001, V = .27	$\chi^2(1, N = 1807) = 88.34$, p <.001, V = .22
B	$\chi^2(1, N = 3741) = 0.03$, $p = .873$, $V = .00$	$\chi^2(1, N = 5178) = 15.58$, p <.001, V = .06	$\chi^2(1, N = 4445) = 33.33$, p <.001, V = .09
C	$\chi^2(1, N = 627) = 0.15$, $p = .693$, $V = .02$	$\chi^2(1, N = 1450) = 54.58$, p <.001, V = .21	$\chi^2(1, N = 1003)$, 36.08, p <.001, V = .19
D	$\chi^2(1, N = 973) = 48.80$, p <.001, $V = .22$	$\chi^2(1, N = 1525) = 84.08$, p <.001, V = .24	$\chi^2(1, N = 954) = 34.65$, p <.001, V = .19

3.6 Discussion

In this study we found that enrolment rates are lower in programs that apply more intensive matching procedures (i.e., more aspects of student-program fit can be tested). This indicates that prospective students are more likely to adjust their initial program choice, if they have the chance to do a more intensive fit-check. Previously, we found that the more matching procedures allow for testing multiple aspects of student-program fit, the more useful prospective students find these for making a final program choice (Soppe et al., 2019). Hence, the types of matching procedures that are deemed useful by prospective students for making a final program choice, are the same types of matching procedures that have lower enrolment rates in matching cohorts and might have lower drop out during the program.

Even though societal factors (e.g., the implementation of academic dismissal policies) cannot be ruled out as possible confounders, and a causal effect cannot be established, it seems probable that these lower enrolment rates in programs with more intensive matching procedures are a result of the implementation of matching procedures, given the differences in enrolment rates between pre-matching cohorts and matching cohorts. Although we observed that the number of students applying for more than one program increased with the introduction of the matching procedures, we showed that this cannot explain the lower enrolment rates in matching cohorts as compared to pre-matching cohorts. As a result, it is plausible that the more intensive matching procedures, being online courses (type 2) and matching days (type 3), help students to check their program-fit and therefore lead to dropout of non-fitting students before the start of the program, and subsequently to lower enrolment rates.

In our analyses we conducted multiple significance tests. This may result in a few false positives as a result of the multiple comparisons problem and, therefore, the results should be interpreted with some caution. However, almost all significant results remain significant after correcting for multiple testing using the Bonferroni method (Beasley & Schumacker, 1995).

This study showed that, overall, enrolment rates were lowest in programs with the most intensive matching procedures (type 3) and highest in programs with the least intensive matching procedures (type 1). There is no indication that there are differences between universities in this overall pattern. However, two deviations from this overall result were found regarding differences between disciplines.

First, investigating differences in enrolment rates across disciplines showed that the programs of discipline C followed a deviating pattern. In this case, enrolment rates are lowest for the program with type 2 procedures (online course) rather than programs with type 3 procedures. We see two plausible explanations for this deviation from the overall pattern. First, the pattern can be caused by an institutional effect. The program of discipline C is the only program of University 3 in our sample. Because the overall enrolment rates of University 3 are unknown, it is not possible to compare type 2 with other matching procedures of this university. Second, discipline C is a STEM program and for students in the STEM domain online matching (type 2) might be more suitable than matching days (type 3) or interviews with students at risk (type 1). The main difference between the online course and matching days is that matching days are on campus, together with other students. This makes it possible for prospective students to test whether they feel a sense of belonging. Sense of belonging might play a different role for STEM students than for non-STEM students when testing their student-program fit (Soppe et al., 2019). Therefore, it would be interesting for future research to explore differences in needs for different kinds of student-program fit testing between students in the STEM domain and students in non-STEM domains.

A second deviation from the overall pattern is the observation that the programs of discipline B show no changes over time in their enrolment rates, unlike the programs of other disciplines. One plausible explanation for this observation could be that it might not be possible to reduce enrolment rates further, given that the enrolment rates in pre-matching cohorts for discipline B were already lower than for other disciplines. Future research could attempt to identify characteristics of programs that already had low enrolment rates before the implementation of matching and investigate what other programs can learn from such characteristics for their own matching procedures.

A potential weakness of our study is the overlap between universities and the type of matching. However, we addressed this potential problem by running the analysis regarding differences in enrolment rates between programs with different types of matching procedures on the data of University 1 separately. The analysis showed no differences in the probability of enrolment between students who participated in an online course or a matching day while this difference was found for the comparison including all universities. This indicates potential differences between universities in the effect of a specific matching procedure on finalizing enrolment. An explanation might be that types of matching that are the same on paper (between universities), are different in practice. Especially an online course or a matching day can be much more intensive at one institution than the other.

This study gives indications that it can be worthwhile to invest in guiding prospective students in their program choice by obliging them to test fit with the program through intensive matching procedures. More intensive activities, in which young people can determine all aspects of fit, lead to lower enrolment rates and thus seem to be more effective than activities that take less effort. Higher education institutions wishing to achieve lower enrolment rates should focus on designing fit checks in such a way that multiple aspects of student-program fit can be tested.

We suggest that these procedures incorporate a test of course material of any kind to provide insight in a student's abilities and an overview of program content, preferably using a representative sample of teaching course materials in any way. Ideally the procedures are on campus, to allow prospective students to interact with one another and the program staff, as well as experiencing the atmosphere in and around the university buildings. The combination of the three components of student-program fit is especially important, since subject-interest is not a stable trait (Vulperhorst et al., 2021) that has been found to decrease in the first year due to feelings of a lack of social integration (Van der Veen et al., 2006). Moreover, students who think they have lower chances of graduating are found to be less integrated, which decreases subject interest even further. By creating a sense of belonging prior to the start of the academic year through on-campus matching procedures, while also dedicating time to verifying ability beliefs and interests, these effects of losing interest could be softened, decreasing the probability that these students will drop out.

Chapter 4

Pre-enrolment Predictors of First-Year Academic Success: Indicators of Student-Program Fit for Different Disciplines

Soppe, K. F. B., Vermue, C.E., Wubbels, T., Klugkist, I.G., & Wijngaards-de Meij, L. D. N. V. (2021, submitted). Pre-Enrolment Predictors of First-Year Academic Success: Indicators of Student-Program Fit for Different Disciplines.

ABSTRACT

Dropout in higher education is a problem because of its negative consequences for both higher education institutions and students. Students experiencing higher levels of fit with their studies are less likely to dropout. Therefore, it would be useful if students could test this fit prior to enrolment. In this paper, we investigated whether such pre-enrolment indicators of fit predict academic success and whether this prediction differs by discipline. Using Structural Equation Modelling, we conducted two separate studies in which we analyzed the same model using data of two Dutch universities. Results indicated that measuring indicators of academic success prior to enrolment is possible. However, these indicators produce small to medium-small effect sizes. Moreover, results show that differential prediction by discipline might help improve predicting first-year academic success. Future studies should account for potential differences between disciplines in their design. The results of this research could help university administrators in improving their intake procedures and, as a result, increase their retention rates.

Author contributions: KS, LWM, TW, & IK designed the study. KS and CV prepared the dataset. KS analyzed the data. KS wrote the paper and CV wrote certain sections of the paper. CV, TW, IK, & LWM provided extensive feedback on the manuscript.

4.1 Introduction

Dropout in higher education is a problem because of its negative consequences for both higher education institutions and students. For higher education institutions high dropout rates go hand in hand with negative outcomes, such as high costs due to (partial) funding based on retention rates (Jongbloed et al., 2018; Kirk, 2018) and complications in enrolment planning (Zajacova et al., 2005). For students, dropping out of higher education is also associated with negative consequences, such as untapped human potential and low return on their financial investment (Jaeger & Page, 1996; Oreopoulos & Petronijevic, 2013; Psacharopoulos, 1994) or reduced social welfare (Hällsten, 2017).

Making a wrong initial program choice is one of the core reasons of dropout worldwide (e.g., Bean, 2005; O'Keefe et al., 2010; Willcoxson & Wynder, 2010; Yorke, 2000). A wrong initial program choice is generally reflected in experiencing feelings of misfit with the program (Feldman et al., 1999; Naylor, et al., 2018; Tinto, 1987; Ulriksen et al., 2010; Warps et al., 2017), that are mainly caused by a realization of a mismatch between expectations and reality of the program content (e.g., Warps et al., 2017, p.11; Watson et al., 2004).

Whereas misfit is associated with a high dropout risk and other negative consequences for both students and higher education institution, good fit between a student and the program is associated with positive outcomes for academic success. For example, students are more engaged in the academic process if they experience a sense of belonging (McFarlane, 2018; Trowler, 2010) and this experience is also important in persisting in university (Kirk, 2018; Tinto, 1987). Believes in one's abilities have also been found to be positively associated with retention (Lent et al., 1984, 1987; Zhang & RiCharde, 1998) as well as with grades (Bong, 2001; Brown et al., 1989; Honicke & Broadbent, 2016; Lent et al., 1984; Multon et al., 1991).

The aim of this study is to identify whether pre-enrolment indicators of fit in intake questionnaires can predict first-year academic success. Non-cognitive factors related to student-program fit, such as interests and self-efficacy, are often measured *during* higher education to identify the relation between actual fit and academic success (Abraham et al., 2012; Robbins et al., 2006). However, the earlier program staff can identify students who might be at risk of dropout the earlier they can offer guidance in the process of adjusting to university life. Moreover, if prospective students can already test their fit with the program before enrolment, they can still opt for another program if they experience a misfit. Therefore, in line with research by

Van Herpen and colleagues (2017) in this study we use *pre-enrolment* aspects of fit as predictors of first-year academic success.

The current study will use data from pre-enrolment questionnaires of two universities in the Netherlands. These questionnaires are a standard element of so-called matching procedures, which are designed as a final check on the program choice prior to enrolment for open-admission programs. Facilitating such a check before enrolment should result in a more fitting student population that is expected to perform better and drop out less. A central aspect of matching procedures at all Dutch universities is the use of intake questionnaires. These questionnaires generally measure aspects such as (a selection of): high school grades; attendance at orientation activities to learn about different programs; motivation or reasons for choosing the program; expectations of the program; expected future jobs; ability beliefs; interests. The questionnaire data, measuring pre-enrolment concepts of fit, will be combined with students' academic records to measure first-year academic success. The focus lies on the first year, because earlier studies have found that the first year denotes students' academic career in later years (e.g., Jansen, 2004; Keup & Stolzenberg, 2004).

This research consists of two separate studies. In Study 1 the relation between pre-enrolment fit indices and first year academic success will be studied for University 1. Study 2 will be a direct replication of the first study, using data of University 2. By conducting separate studies, we aim to investigate whether the relation between pre-enrolment fit indices and first year academic success holds in different contexts.

4.2 Pre-Enrolment Predictors of Academic Success

Academic success relies on general academic ability, such as reading, writing and math skills, and on motivation as the drive to persist (Campbell, 1990). Ability has been a major component of many theories on academic success and retention (e.g., Bean, 2005; Tinto, 1975) and is widely used as a selection requirement in admission procedures (Harackiewicz et al., 2002). Ability is generally measured through standardized test scores (such as SAT and ACT) or high school grades. Like ability, motivation is a prominent component in student retention research (e.g., Deci & Ryan, 2008). Motivation is the drive that keeps students going, after the initial stage of learning a new skill (Zyphur et al., 2007). This drive is reflected in the personality trait conscientiousness, as it involves the ability to adapt to goals set by others (Hough & Schneider, 1996) and the tendency to carry out tasks in a careful, dutiful and organized manner (Goldberg, 1990; Tross et al., 2000). It has been suggested that the adaptability component of conscientiousness reflects general ability, whereas the achievement striving and dutifulness aspects are motivational components (Alarcon & Edwards, 2013). In this study we will use high school performance and conscientiousness as indicators of general academic ability and motivation.

To succeed in higher education students require more than a general ability and motivation to study. Students would need to fit with the program of their choice, to remain motivated to study. Person-environment fit is seen as the compatibility between individual and environmental characteristics (Kristof-Brown & Guay, 2011). We argue that to be able to test fit, prospective students should be able to: 1) verify the accuracy of their ability believes (i.e., does the extent to which they believe they are capable of performing a task, match with reality (Bandura, 1977; Deci & Ryan, 2008)); 2) identify to what extent they value the content (e.g., do their interests match with the content of the program (Wigfield, 1994; Wigfield & Eccles, 1992)); and 3) experience the social aspects of the new environment (Tinto, 1987), because the type of students and staff involved is also an inherent part of identifying with and belonging to a higher education program (Hoffman et al., 2002; Swenson et al., 2008). The third aspect of fit cannot be assessed very well through a questionnaire and will therefore not be included in this study. Lastly, not only whether prospective students expect to fit with the program, but also the amount of effort they put into being able to assess this fit is relevant in this regard. In other words, students who have explored their different options more elaborately, might be more capable to choose the best fitting program and thus extensively exploring one's options might positively influence their academic success. Therefore, we include participation in orientation activities in our study.

This paper does not only want to contribute to knowledge about pre-enrolment measures of fit, but also to the systematic theoretically grounded use of Structural Equation Modeling (SEM) in Higher Education research. In a methodological review on all SEM articles published in top tier higher education journals up until 2013, Green (2016) showed that SEM is inconsistently applied within and across journals and that misunderstandings in SEM's application appear to be widespread. One of the biggest issues Green identified, is a lack of building theoretically supported models (Green 2016, p. 2132 / p. 2143). This study will contribute to the SEM literature in the field of higher education research by giving thorough attention to a sound theoretically grounded model. The theoretical model of this study is shown in Figure 4.1 and the concepts are discussed in more detail below.

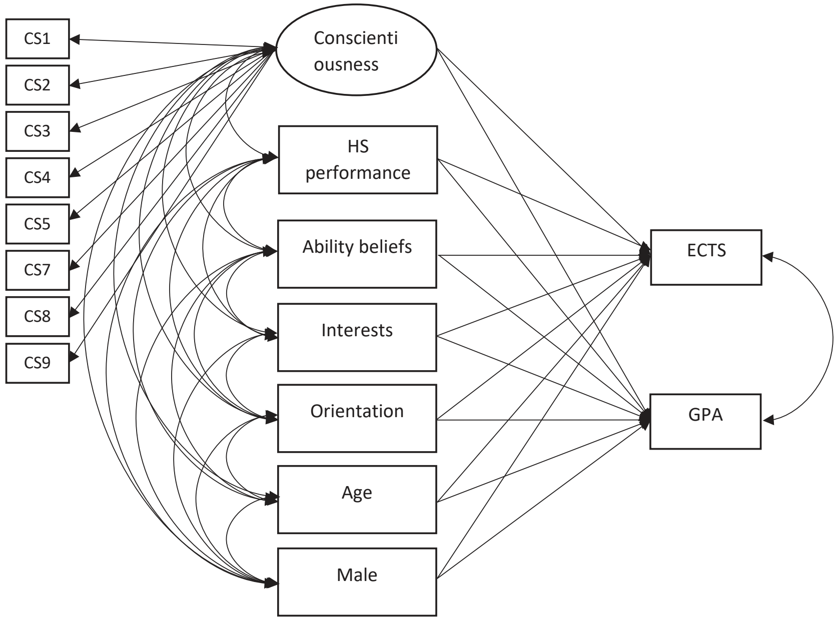

Figure 4.1 Theoretical model.

4.2.1 Conscientiousness

Conscientiousness, one of the personality traits of the Big Five Inventory (John & Srivastava, 1999), is commonly associated with high levels of self-control, dependability, persistence, orderliness and the tendency to work hard and diligently (e.g., Busato et al., 2000; Eisenberg et al., 2014). Moreover, in learning contexts, conscientiousness is associated with employing effective learning strategies (Komarraju et al., 2011), which results in motivation to keep students going (Alarcon & Edwards, 2013). Conscientiousness has often been found to be an important predictor of academic success (e.g., Busato et al., 2000; De Fruyt & Mervielde, 1996; Kappe & Van der Vlier, 2010; Poropat, 2009). In fact, research showed that it is the personality trait that most strongly predicts academic success (e.g., O'Connor & Paunonen, 2007; Trapmann et al., 2007). Because of its high predictive validity, regardless of which academic performance measure is used, conscientiousness has been suggested as a useful criterion in admission and selection processes in higher education (Conard, 2006; Kappe & Van der Flier, 2010, 2012; Kling et al., 2012).

4.2.2 High School Performance

Students' general abilities are also important for their academic success. In this study, high school grades on Dutch, English and Mathematics will be used as a proxy for High School GPA (HSGPA). Many researchers found HSGPA to be the strongest predictor of academic success (e.g., Camara & Echternacht, 2000; Munro, 1981; Tross et al., 2000; Zheng et al., 2002). However, more recently the use of HSGPA in selection procedures is debated (Niessen, 2017; Steenman et al., 2016; Vulperhorst et al., 2017). HSGPA is, however, often the only available measure for prior achievement (Pitman, 2016). Although usually treated as such, HSGPA is not purely a cognitive predictor of academic success (Niessen & Meijer, 2020). A substantial amount of the variance in HSGPA can be explained by non-cognitive traits, such as conscientiousness and self-efficacy (Borghans et al., 2011; Deary et al., 2007; Dumfart & Neubauer, 2016).

Regardless the discussion on using HSGPA for selection purposes, HSGPA remains an important predictor for retention (Westrick et al., 2015), as well as first-year GPA (FYGPA). For this reason, Zwick (2017) argues that high school grades should play an important role in higher education admissions.

4.2.3 Ability Beliefs

The concept ability beliefs is an overarching term that captures several sub-concepts regarding beliefs about one's capabilities, such as perceptions of competence and perceptions of the difficulty of different tasks (Eccles & Wigfield, 2002). Eccles and colleagues initially distinguished between beliefs about ability, i.e., an evaluation of one's competence in different areas, and expectancies for success, i.e., beliefs regarding performance on upcoming tasks (Eccles et al., 1983). However, their empirical research has shown that adolescents do not distinguish between these two levels of beliefs in real-world achievement situations (Eccles & Wigfield, 2002). The part of Eccles' ability beliefs concept regarding expectancies for success is closely related to Bandura's (1997) well-known concept of self-efficacy, otherwise known as personal efficacy expectations. Just like the concepts *expectancy beliefs* (Eccles et al., 1983) and *competences* (Deci & Ryan, 2008), self-efficacy is seen as a motivational belief. Within the context of higher education research, self-efficacy is often applied to the educational context by specifically measuring *academic* self-efficacy, i.e., students' understanding of their ability to learn and perform (Bandura, 1997; Schunk & Pajares, 2009). These ability beliefs result from assessing one's own as well as other's past performance, received feedback on capabilities and performances and feelings regarding these capabilities and performances.

Self-efficacy has been shown to be a powerful predictor of motivation and achievement in higher education (see for example reviews by Abraham et al., 2012 and Brown et al., 2008). High levels of self-efficacy are related to both high achievement and retention (Pajares, 1996, 1997; Schunk, 1995; Zajacova et al., 2005). These positive relations are found in studies measuring academic self-efficacy during higher education. Very little is known, however, about the relation between academic self-efficacy and academic success during the transition to higher education (Van Herpen et al., 2017). Although it can be expected that prospective students are able to assess their academic abilities based on past experiences in high school (Pintrich, 2004; Schunk & Pajares, 2009), Van Herpen and colleagues (2017) found no relationship between pre-university academic self-efficacy and obtained number of credits in the first year. In this study we will explore the relation of the broader concept pre-university ability beliefs to credits, as well as to FYGPA.

4.2.4 Interests

Vocational interest plays an important role in adolescents' learning and development (Renninger & Hidi, 2017), and is related to program choice (Whitney, 1969). However, it may be difficult to decide which interests to pursue. Students often have multiple interests, but they cannot pursue them all in one higher education program (Hofer, 2010; Vulperhorst et al., 2018). Hence, they will have to weigh these multiple interests, trying to maximize pros and minimize cons (Eccles & Wigfield, 2002).

Experiencing interest is linked to high motivation to learn (Harackiewicz et al., 2008; Hidi & Renninger, 2006), and high levels of intrinsic motivation, in return, result in higher levels of academic success (Deci & Ryan, 2008). Students who chose not to pursue a specific program or interest may come to regret this (Kucel & Vilalta-Bufi, 2013). These feelings of regret may result in dropout (Holmegaard et al., 2016). Moreover, Schelfhout and colleagues (2019) found that program interest is also directly related to academic performance.

4.2.5 Program Orientation

In the Netherlands, prospective students have extensive possibilities to explore which higher education program suits them best. For each phase in their process of choosing a program, there are sources of information or activities available that should guide them in making a fitting choice. These orientation activities aid prospective students in the transition from a high school to university environment (Fitz-Walter et al., 2014). In the early stages of their program choice, prospective students can, for example, search university websites to get an idea of the academic discipline that would suit them or visit *open days* at which there are sessions regarding the institution in general as well as mini-lectures for specific programs. Once prospective students have narrowed down their choice to several programs, there are possibilities to attend so-called *student-for-a-day* activities or talk to study advisors of specific programs. These latter activities are partially aimed at making students feel part of the community and are meant to create a general sense of belonging with university life and the academic community before prospective students finalize their enrolment. It is known that orientation weeks for new students (Murphy et al., 2002; Nguyen et al., 2018) as well as new employees in companies (Acevedo & Yancey, 2011) have a significant impact on future learning efficiency. We argue that the same may apply to these pre-university orientation activities; making use of these activities, will enhance student-program fit. This should, in return, result in a higher performance.

4.2.6 Age and Gender

In the Netherlands, age at university-entry is strongly related to a student's route to higher education. Students who directly transition from high school (pre-university level) are on average younger than those who transition to university from a university of applied sciences. Older students are expected to possess better coping strategies and higher self-esteem, which allows them to adapt better to university situations (Clifton et al., 2008). However, mixed findings are reported regarding age and academic success. Older students are found to achieve higher grades (Clifton et al., 2008; Etcheverry et al., 2001; Sheard, 2009), but also to dropout more often (Crosta, 2013; Lassibille & Navarro Gómez, 2008). Other studies found no association between age and academic achievement (Farsides & Woodfield, 2007; Ting & Robinson, 1998). Meta-analytic results from Richardson and colleagues (2012) showed that age and other demographics (i.e., gender and socio-economic status) matter, but that effect sizes are small.

Even though general cognitive ability is similar across gender (Carvalho, 2016; Pintrich & de Groot, 1990), female students outperform their male peers by obtaining higher grades and more credits (e.g., Conger & Long, 2010), as well as being more likely to graduate (e.g., Conger & Long, 2010; Hofman & Van den Berg, 2000). However, in a study on older students, Gigliotti and Huff (1995) found that gender influences GPA, but not retention in year 2.

4.3 Differential Prediction by Discipline

The predictive power of indicators of academic success is often studied for large, university-wide samples (Beaulac & Rosenthal, 2019) or specific disciplines, such as psychology (e.g., Busato et al., 2000) or STEM (science, technology, engineering, and mathematics) programs (e.g., Holmegaard et al., 2016). Little is known, however, about the relevance of discipline in this regard. Certain indicators might be more important predictors for academic success for certain groups of students.

Fonteyne and colleagues (2017) tested the relation between cognitive and non-cognitive predictors and academic achievement. Their results showed that the inclusion of non-cognitive factors allows for better prediction of academic achievement for some programs but not for others. Moreover, they showed that the predictive power of variables varied across a range of different study programs, suggesting that research findings about the prediction of academic achievement might benefit from considering the specific program context. A range of programs were included in that study, but STEM programs were not. Despite their call for research that includes STEM programs, no studies have been found. However, Willems and colleagues (2019) found that, after controlling for cognitive factors, non-cognitive factors did not contribute to the prediction of academic achievement for STEM students. In this study, we'll explore for both cognitive and non-cognitive factors whether their predictive power differs between STEM students and other university students in our sample.

4.4 Research Questions and Hypotheses

Our study evolves around several research questions. The first question concerns the absolute and relative effects of conscientiousness, high school performance, ability beliefs, interests, extensiveness of program orientation, age and gender on academic achievement. We hypothesize that conscientiousness (e.g., Busato et al., 2000), high school performance (e.g., Camara & Echternacht, 2000), ability beliefs (e.g., Brown et al., 2008), interests (e.g., Hidi & Renninger, 2006), and extensiveness of program orientation (e.g., Nguyen et al., 2018) will be positively related to both credits (H1) and FYGPA (H2). Moreover, we expect medium effects for conscientiousness and high school performance on both outcome measures (H3) and small effects for ability beliefs, interests and extensiveness of program orientation on both outcomes (H4). Furthermore, we expect the effect of age on FYGPA (e.g., Clifton et al., 2008) to be positive (H5a) and the effect of age on credits (e.g., Crosta, 2013) to be negative (H5b). Lastly, we expect female students to outperform male students regarding both outcome measures (e.g., Conger & Long, 2010) (H6a) and the effect of gender on FYGPA to be stronger than the effect of gender on credits (e.g., Gigliotti & Huff, 1995) (H6b).

Second, we explore whether the relation between the predictors and outcomes differ between disciplines. Based on the small amount of research on this topic, it can be hypothesized that the relation between especially non-cognitive factors and academic success differs between STEM students and non-STEM students (Fonteyne et al., 2017; Willems et al., 2019). Hence, we expect that the relation between conscientiousness, ability beliefs, interests, extensiveness of orientation and academic achievement differs between STEM and non-STEM studies (H7), while the relation between high school performance and academic achievement may not (H8).

4.5 Study 1

4.5.1 Method

4.5.1.1 Sample

The sample consists of 10,151 students (49.1% female) who enrolled in a open-admission bachelor program of University 1 in the cohorts 2014 to 2016. University 1 is one of the eight publicly funded general research universities in the Netherlands. Students were on average almost 20 years old (*M* = 19.65; *SD* = 2.70) at the start of the academic year. 30.7% of the students were enrolled in a STEM program.

4.5.1.2 Procedure and questionnaires

Permission for this study was obtained from the Ethical Review Board of one of the universities in this study (review number FETC20-080).

At University 1, the university-wide intake questionnaire was developed by experts from Teacher Education, the Faculty of Economics and Business and the department of Educational Support and Innovation. Besides questions concerning background information, information about prior education and the attendance of orientation activities, four major topics were included based on literature (Guiffrida et al., 2013; John & Srivastava, 1999; Krause & Coates, 2008; Pampaka et al., 2012): quality of the student experience, being able to handle expectations, academic integration, and fit with the program. From these overarching topics, the scales regarding motivation and fit were used in the present study.

4.5.1.3 Measures

Academic achievement. Academic achievement is operationalized as both credits and grades. Credits were measured by the total amount of ECTS (European Credit Transfer System) obtained in the program a student was enrolled in. Students need to obtain 45 out of 60 ECTS or more to re-enroll in their current program for their sophomore year. FYGPA is the average grade of all courses that a student participated in, in their first year of the program, ranging from 1-10.

Conscientiousness. Students' self-perceived conscientiousness was measured using nine items. These items were taken from the Dutch translation of the Big Five Inventory (Denissen et al., 2008). Each item was answered on a 5-point Likert scale ranging from (1) strongly disagree to (5) strongly agree. Negatively phrased items were recoded so that higher scores indicated higher levels of self-perceived conscientiousness. One item was deleted after performing a reliability analysis (see Appendix A.I), resulting in a Cronbach's alpha of .78 for the eight remaining items. This indicates good internal consistency (Kline, 2011).

High school performance. To measure high school performance (on a scale from 1-10), a mean of the three core subjects in the pre-final year of high school as self-reported in the questionnaire was taken (Dutch, English and Mathematics). If one of the grades was missing, a mean of the remaining two subjects was taken. When grades for only one subject were available, the variable was coded as missing. This measure was only available for students with a pre-university high school degree (VWO).

Ability beliefs. Students' belief in their own abilities was measured using a single item from the questionnaire, stating: "This program matches my skills and talent" ranging from 1 (totally disagree) to 7 (totally agree).

Interests. Students' interest in the program was measured using the statement "I am interested in this field". Answers could range from 1 (totally disagree) to 7 (totally agree).

Program orientation. The extent to which a student participated in activities to explore whether the program of their choice suits them, was measured by taking the sum of 8 dichotomous items. Example activities are "participating in a university open day" or "talking to a study advisor". Each item could be answered with (1) *yes* or (0) *no*, resulting in a variable ranging from 0-8.

Gender. Information about students' gender was taken from the university registration systems. Male students were coded as 1 and female students as 0.

Age. Students' date of birth was taken from the university registration systems and then recoded into age at the start of the program by subtracting date of birth from the first day of the academic year for each cohort.

4.5.2 Analyses

In this study we use Structural Equation Modelling (SEM) for testing structural models using latent variables (Bollen, 1989). The main advantage of SEM over regression is the possibility to combine latent unobserved variables with observed variables in one predictive model (Keith, 2006).

To test the hypothesized model, MPlus, version 8.3 (Muthén & Muthén, 2019) was used. First, to determine whether data of the separate cohorts could be combined, measurement invariance by cohort was established for the latent variable conscientiousness. After establishing configural invariance, a model test was performed comparing a configural, metric (equal factor loadings), and scalar invariant model (equal factor loadings and intercepts) to one another. Partial scalar invariance was achieved. Results can be found in Appendix A.II. Hence, the decision was taken to combine the cohorts. Second, the hypothesized model (Figure 4.1) was tested, without taking discipline into account.

Third, before examining whether the structural model differed for STEM students compared to non-STEM students, partial scalar invariance was established for the measurement model (see Appendix A.III). Since full scalar invariance could not be established, a sensitivity analysis was conducted to establish which parameters in the model were affected by this partial measurement non-invariance. Results of the sensitivity analysis showed that all parameters were stable across the different levels of invariance, indicating that none of the parameters were affected by the partial non-invariance. Hence, we proceeded with the least complex model, the one with no invariance. Finally, using multi-group analysis, the structural model was specified for STEM students and non-STEM students separately.

4.6 Results

4.6.1 Descriptive Statistics

Table 4.1 shows descriptive statistics for the total sample and for STEM students and non-STEM students separately.

Table 4.1 Descriptive statistics for University 1 for STEM and non-STEM students.

University 1	STEM	non-STEM	Total
	M (SD)	M (SD)	M (SD)
Conscientiousness	3.63 (0.53)	3.67 (0.53)	3.66 (0.53)
Ability beliefs	5.93 (0.78)	6.01 (0.84)	5.98 (0.82)
Interests	6.55 (0.57)	6.57 (0.59)	6.56 (0.58)
Program orientation	5.36 (1.29)	5.31 (1.37)	5.32 (1.35)
High school performance	6.73 (0.65)	6.64 (0.59)	6.67 (0.60)
Age	19.18 (1.64)	19.86 (3.03)	19.65 (2.70)
Male	0.61 (0.49)	0.47 (0.50)	0.51 (0.50)
Credits	45.35 (18.73)	43.68 (20.43)	44.19 (19.93)
FYGPA	6.85 (1.15)	6.59 (1.17)	6.67 (1.17)
Range of *n*	2848-3113	5743-7038	8695-10151

4.6.2 Structural Model

First, the relation between the predictors and academic success was tested for all students together. Since the model fit was not directly adequate, 3 item correlations between the error terms of indicators of conscientiousness were added, based on Modification Indices and theoretical considerations (see appendix A.II and A.III). The fit of the modified model was acceptable; CFI = .936, RMSEA = .046.

Table 4.2 presents the standardized parameter estimates of the model. Given the large sample size, an alpha-level of 1% was used for significance testing. The model as a whole explained 7.6% of the variance in credits and 13.5% of the variance in FYGPA. The effects of conscientiousness, high school performance, interests, and extensiveness of program orientation on both credits and FYGPA are positive and significant. However, the effects of ability beliefs on both outcomes are not significant, leading us to partially accept H1 and H2. As expected from the literature, the largest effects are found for conscientiousness and high school performance. However, the effect sizes can be interpreted as small (Cohen, 1988) rather than medium, leading us to reject H3. The effects of interests and extensiveness of orientation on both outcomes can indeed be interpreted as small. But since the effects of ability beliefs are not significant, H4 can only be partially accepted. In line with our hypothesis, the effect of age on credits is negative. However, contrary to our hypothesis, the effect of age on FYGPA is negative as well, leading us to accept H5a while rejecting H5b. Lastly, as hypothesized, male students obtain less credits and a lower FYGPA than female students, leading us to accept H6a. However, the effect of gender on FYGPA is not stronger than the effect of gender on credits, resulting in the rejection of H6b.

4.6.3 Differences by Discipline

First, we tested the fit of the full structural model without equality constraints on the parameters across STEM and non-STEM students. This model had an acceptable fit; CFI = .935, RMSEA = .045. To identify whether the relation between each of the parameters differs for STEM and non-STEM students, we tested whether the paths between the predictors and the outcome variables could be constrained to be equal for both groups. The fully constrained model still had acceptable fit; CFI = .932, RMSEA = .044. However, according to our main decision rule (ΔCFI $\leq$0.002), the fully constrained model fits the data significantly worse than the unconstrained model, indicating that at least one of the parameters in the model could not be restricted to be equal across disciplines.

Since Chi-square difference testing is not adequate given our large sample size, differences between STEM and non-STEM students are assessed based on the differences between the standardized effects. A ratio of 1.5 is used to determine whether there is a substantial difference between STEM and non-STEM students. In other words, if the effect of one group on one of the outcome variables is 1.5 times larger than the effect of the other group, the difference is deemed to be substantial. Standardized parameter estimates of the model for STEM and non-STEM students separately are included in Table 4.2.

Table 4.2 Standardized effects of pre-enrolment predictors on credits and FYGPA per university for the sample and split by STEM vs. non-STEM.

	UNIVERSITY 1						UNIVERSITY 2					
	Credits[8]			FYGPA			Credits[9]			FYGPA		
	Total	STEM	non-STEM	Total	STEM	non-STEM	Total	STEM	non-STEM	Total	STEM	non-STEM
Conscien-tiousness	.076 (.013) *<.001*	.043 (.024) *.069*	.091 (.016) *<.001*	.099 (.013) *<.001*	.044 (.025) *.075*	.129 (.016) *<.001*	.140 (.010) *<.001*	.085 (.024) *<.001*	.154 (.010) *<.001*	.175 (.009) *<.001*	.148 (.022) *<.001*	.186 (.010) *<.001*
High school performance	.127 (.010) *<.001*	.121 (.018) *<.001*	.122 (.012) *<.001*	.286 (.011) *<.001*	.322 (.020) *<.001*	.253 (.014) *<.001*	.207 (.008) *<.001*	.313 (.019) *<.001*	.173 (.009) *<.001*	.366 (.008) *<.001*	.544 (.016) *<.001*	.320 (.009) *<.001*
Ability beliefs	-.020 (.011) *.059*	-.002 (.019) *.930*	-.025 (.013) *.048*	-.005 (.011) *.664*	.047 (.019) *.016*	-.022 (.013) *.083*	-.026 (.008) *.002*	-.037 (.019) *.052*	-.031 (.009) *.001*	.005 (.008) *.487*	.039 (.016) *.016*	-.013 (.008) *.130*
Interests	.047 (.011) *<.001*	.040 (.019) *.036*	.047 (.013) *<.001*	.040 (.010) *<.001*	.035 (.018) *.054*	.037 (.013) *.003*	.049 (.009) *<.001*	-.013 (.022) *.576*	.053 (.009) *.001*	.142 (.010) *<.001*	.028 (.017) *.094*	.143 (.010) *<.001*
program Orientation	.078 (.010) *<.001*	.100 (.019) *<.001*	.067 (.012) *<.001*	.046 (.010) *<.001*	.059 (.019) *.002*	.037 (.012) *.003*	-.016 (.008) *.045*	.044 (.020) *.029*	-.028 (.008) *.001*	-.024 (.007) *.001*	.053 (.018) *.003*	-.021 (.008) *.009*
Age	-.141 (.010) *<.001*	-.200 (.017) *<.001*	-.132 (.012) *<.001*	-.094 (.017) *<.001*	-.132 (.019) *<.001*	-.078 (.020) *<.001*	-.079 (.009) *<.001*	-.114 (.030) *<.001*	-.075 (.010) *<.001*	-.032 (.009) *<.001*	-.026 (.016) *.106*	-.026 (.010) *.011*
Male	-.053 (.011) *<.001*	-.073 (.020) *<.001*	-.047 (.013) *<.001*	-.040 (.010) *<.001*	-.049 (.018) *.008*	-.058 (.013) *<.001*	-.047 (.008) *<.001*	.067 (.020) *.001*	-.071 (.009) *<.001*	-.040 (.008) *<.001*	.075 (.018) *<.001*	-.078 (.009) *.011*
explained variance	7.6%	10.5%	7.0%	13.5%	17.2%	12.2%	9.3%	14.4%	8.9%	23.6%	38.8%	21.6%

Note: Table contains per cell Standardized effect, Standard Error (between brackets), and p-value (italic).

[8] The correlations between credits and FYGPA for the respective models were .776 (total model), .765 (STEM model), and .785 (non-STEM model).

[9] The correlations between credits and FYGPA for the respective models were .603 (total model), .482 (STEM model), and .614 (non-STEM model).

The multigroup model explained 7.0% of the variance in credits for non-STEM students and 10.5% of the variance in credits for STEM students as well as 12.2% of the variance in FYGPA for non-STEM students and 17.2% of the variance in FYGPA for STEM students. Only indicators for which the ratio between STEM and non-STEM is 1.5 or larger are discussed. For students in STEM programs conscientiousness does not influence the obtained number of credits nor FYGPA. In contrast, the positive effect of conscientiousness on both credits and FYGPA is significant for non-STEM students. In other words, more conscientious students in non-STEM programs obtain more credits and a higher FYGPA than their less conscientious peers. The positive effect of orientation on both credits and FYGPA is stronger for STEM than for non-STEM students. That is, students who have taken part in more orientation activities obtain more credits and a higher FYGPA and this effect is stronger for STEM students. For the relations of age on both credits and FYGPA, the negative effect is stronger for STEM students. Put differently, older students obtain less credits and a lower FYGPA than younger students and these effects are stronger for STEM students than for non-STEM students. Lastly, female students obtain more credits than male students, and this effect is stronger for STEM students than for non-STEM students.

Since most, but not all the non-cognitive predictors differ between STEM and non-STEM students, H7 can be partially accepted. Moreover, H8 can be accepted, because the effect of high school performance on both outcome variables does not differ substantially between STEM and non-STEM students.

4.7 Study 2

4.7.1 Method

4.7.1.1 Sample

The sample consists of 15,460 students (52.1% female) who enrolled in an open-admission bachelor program at University 2 in 2015 and 2016. University 2 is one of the eight publicly funded general research universities in the Netherlands. Participants were on average 19 years old (M = 19.08; SD = 2.63) at the start of the program. 15.8% of the students were enrolled in a STEM program.

4.7.1.2 Procedure and questionnaires

At University 2 matching was introduced in 2012 as a pilot in several programs. Matching consisted of filling out an online questionnaire, followed by an on-campus activity. After the pilot year, matching was introduced for the whole university in 2013. The questionnaire was improved each year, until it took its final shape in

2015. The distribution of the questionnaire is automated; as soon as students file an admission request to a program, they receive the questionnaire by email. Upon return of a completed questionnaire, students receive an invitation to take part in an on-campus activity, which is compulsory to successfully complete the matching procedure.

The questionnaire was developed in a two-fold approach. General questions concerning background information, information about prior education and the attendance of orientation activities as well as questions concerning fit with higher education in general were developed by a university-wide task force. On top of those central questions, staff members of each program could add program-specific questions to the questionnaire. As a result, questionnaires are tailored to each program, while at the same time they all contain the same questions about motivation, ability and fit.

4.7.1.3 Measures

Academic achievement, conscientiousness, gender, and age were operationalized in the exact same manner as for University 1. For conscientiousness, the same item as for University 1 was deleted after performing a reliability analysis (see Appendix B.I), resulting in a Cronbach's alpha of .80 for the eight remaining items. This indicates very good internal consistency (Kline, 2011).

Ability beliefs. Students' belief in their own abilities was measured using a single item from the questionnaire, stating: "the program matches my capacities and skills". The item could be answered with (1) yes or (0) no.

Interests. Students' interest in the program was measured using the statement "the program is in line with my interests". The item could be answered with (1) yes or (0) no.

Program orientation. The extensiveness of program orientation is operationalized in the same manner as for University 1. Exact wording of the phrases differed slightly between the universities, but the questions concerned the same activities.

High school performance. High school performance was operationalized in the same manner as for University 1. However, for University 2, it was not just available for students with a pre-university degree but also for almost all other students. For students transitioning from a university of applied sciences, grades from their final year were used, for all others, just like in University 1, grades from the pre-final high school year were used.

4.7.2 *Analyses*

First, measurement invariance by cohort was tested in the same manner as for University 1. For University 2, full scalar invariance was achieved. Results can be found in Appendix B.II. Hence, the decision was taken to combine the cohorts. Second, the hypothesized model (Figure 4.1) was tested, without taking discipline into account. Third, before examining whether the structural model differed for STEM students compared to non-STEM students, partial scalar invariance was established for the measurement model (see Appendix B.III). Since full scalar invariance could not be established, a sensitivity analysis was conducted to establish which parameters in the model were affected by this partial measurement non-invariance. Results of the sensitivity analysis showed that all parameters were stable across the different levels of invariance, indicating that none of the parameters were affected by the partial non-invariance. Hence, we proceeded with the least complex model, the one with no invariance. Finally, using multi-group analysis, the structural model was specified for STEM students and non-STEM students separately.

4.8 Results

4.8.1 *Descriptive Statistics*

Table 4.3 shows descriptive statistics for the total sample and for STEM students and non-STEM students separately.

Table 4.3 Descriptive statistics for University 2 for STEM and non-STEM students.

University 2	STEM	non-STEM	Total
	M (SD)	M (SD)	M (SD)
Conscientiousness	3.68	3.74	3.73
	(0.56)	(0.55)	(0.55)
Ability beliefs	0.79	0.64	0.67
	(0.41)	(0.48)	(0.47)
Interests	0.98	0.91	0.92
	(0.14)	(0.29)	(0.28)
Program orientation	3.02	3.71	3.60
	(1.48)	(1.53)	(1.54)
High school performance	6.98	6.80	6.83
	(0.75)	(0.59)	(0.61)
Age	18.77	19.12	19.07
	(2.38)	(2.75)	(2.70)
Male	0.66	0.43	0.46
	(0.47)	(0.49)	(0.50)
Credits	42.17	41.22	41.36
	(21.33)	(19.74)	(19.99)
FYGPA	7.11	6.66	6.73
	(0.76)	(1.00)	(0.98)
Range of *n*	2112-2342	11782-13118	13820-15460

4.8.2 *Structural Model*

First, the relation between the predictors and academic success was tested for all students together. Since the model fit was not adequate, 3 item correlations between the error terms of indicators of conscientiousness were added, based on Modification Indices and theoretical considerations (see appendix B.II and B.III). The fit of the modified model was acceptable; CFI = .939, RMSEA = .046.

Table 4.2 presents the standardized parameter estimates of the model. The model as a whole explained 9.4% of the variance in credits and 23.6% of the variance in FYGPA. The effects of conscientiousness, high school performance, and interests on both credits and FYGPA are positive and significant. However, the effects of ability beliefs on credits and orientation on both outcomes are negative, and the effect of ability beliefs on FYGPA is not significant, leading us to partially accept H1 and H2. In line with our hypotheses, the largest effects are conscientiousness and high school performance. However, the effect sizes can be interpreted as small to medium-small (Cohen, 1988) rather than medium, leading us to reject H3. The effects of interests on both outcomes can indeed be interpreted as small. However, since the effects of ability beliefs and orientation are negative or not significant, H4 can only be partially accepted. In line with our hypothesis, the effect of age on credits is negative, but the effect of age on FYGPA is negative as well, leading us to accept H5a while rejecting H5b. Lastly, as hypothesized, male students obtain less credits and a lower FYGPA than female students, leading us to accept H6a. However, the effect of gender on FYGPA is not stronger than the effect of gender on credits, resulting in the rejection of H6b.

4.8.3 *Differences by Discipline*

First, we tested the fit of the full structural model without equality constraints on the parameters across STEM and non-STEM students. This model had an acceptable fit; CFI = .933, RMSEA = .047. To identify whether the relation between each of the parameters differs for STEM and non-STEM students, we tested whether the paths between the predictors and the outcome variables could be constrained to be equal for both groups. The fully constrained model still had acceptable fit; CFI = .930, RMSEA = .046. However, according to our main decision rule (ΔCFI ≤0.002), the fully constrained model fits the data significantly worse than the unconstrained model, indicating that at least one of the parameters in the model could not be restricted to be equal across disciplines. Standardized parameter estimates of the model for STEM and non-STEM students separately are included in Table 4.2.

The multigroup model explained 8.9% of the variance in credits for non-STEM students and 14.4% of the variance in credits for STEM students as well as 21.6% of the variance in FYGPA for non-STEM students and 38.8% of variance in FYGPA for STEM students. Only indicators for which the ratio between STEM and non-STEM is 1.5 or larger are discussed. The positive effect of conscientiousness on credits is stronger for non-STEM students. In other words, more conscientious students obtain more credits than their less conscientious peers and this effect is stronger for non-STEM students. The positive effect of high school performance on both credits and FYGPA is stronger for STEM students. This indicates that students with better high school grades obtain more credits and a higher FYGPA than students with lower high school grades and that this effect is stronger for STEM students. For interests, the effect on both credits and FYGPA is only significant for non-STEM students. Students in non-STEM programs with stronger interests in the program obtain more credits, while for students in STEM programs there is no relation between their interests and obtained number of credits. The effect of orientation on both credits and FYGPA is positive for STEM students and negative for non-STEM students. That is, students in STEM programs who have taken part in more orientation activities obtain more credits and a higher FYGPA, while students in non-STEM programs who have taken part in more orientation activities obtain less credits and a lower FYGPA. Notably, the positive effect on credits for STEM students in nonsignificant at our chosen alpha-level of 1%. The negative effect of age on credits is stronger for STEM students than for non-STEM students. Put differently, older students obtain less credits and this effect is stronger for STEM students. Lastly, the effect of gender on both credits and FYGPA is positive for STEM students and negative for non-STEM students. That is, in STEM programs male students obtain more credits and a higher FYGPA, while in non-STEM programs female students obtain more credits and a higher FYGPA. Notably, the negative effect of gender on FYGPA for non-STEM students is nonsignificant at our chosen alpha level of 1%.

Since most, but not all the non-cognitive predictors differ between non-STEM and STEM students, H7 can be partially accepted. Moreover, H8 must be rejected, because the effect of high school performance on both outcome variables differs between STEM and non-STEM students.

4.9 General Discussion

The objective of this study was to investigate pre-enrolment predictors of first-year academic success in two Dutch universities. In this section we summarize and discuss the results for both universities, discuss implications for research and practice, report limitations of this study and provide recommendations for future research.

An important finding of this study is that indicators of academic success, as measured prior to enrolment, are predictive of academic success with small to medium-small effect sizes. Especially non-cognitive factors, such as motivation and effort, are often measured *during* higher education, to predict first-year academic success (Abraham et al., 2012; Robbins et al., 2006). However, this study has shown that non-cognitive factors measured prior to enrolment, such as conscientiousness and interests, add to the prediction of first-year academic success. The fact that we saw almost identical patterns across the two universities when estimating the model as a whole, strengthens these findings further.

A second important finding of this study is that it is important to distinguish between different disciplines when researching academic success. We have shown that for most of our indicators, there is a difference in the strength and/or direction of the effects between STEM and non-STEM students. Interests, for example, positively relate to both credits and FYGPA in both universities for non-STEM students. However, for STEM students there is no relation between interests and either of the outcome measures in both universities. Also conscientiousness appears to be a more important predictor of academic success for non-STEM students than for STEM students. In contrast, high school performance, number of attended orientation activities, and age seem to be more important in predicting academic success for STEM students than for non-STEM students. This may indicate that non-cognitive traits are stronger predictors of academic success for non-STEM students, while the high school performance is a stronger predictor of academic success for STEM students.

Lastly, when looking at the main effects, it appears that in both universities, female students obtain more credits and a higher FYGPA than male students. However, when inspecting the effects for STEM and non-STEM programs separately, it becomes clear that in University 2 male students in STEM programs outperform their female peers regarding both credits and FYGPA. On the other hand, in University 1 female students obtain more credits and a higher FYGPA in both the STEM and the non-STEM programs. One explanation for this difference might be that both universities do not offer exactly the same programs. It is possible that the STEM programs at University 2 have, on average, a stronger STEM profile. That would explain the underperformance of female students, since many studies (e.g., Wang & Degol, 2017) have shown the persistent underrepresentation in, as well as underperformance of female students in STEM programs.

4.9.1 Limitations and Directions for Future Research

One of the main aims of this research was to test the external validity of our model by applying it to two different settings. To do so, the data for both universities had to be made as comparable as possible. The biggest difference in the available data concerned the concepts ability believes and interests. University 1 surveyed these items with an intrinsic motivation scale (Warps et al., 2009) with low reliability (α = .64). Therefore, we matched the single-item concepts of University 2 to the best-fitting items of the scale used by University 1. Moreover, for University 1 the questions about ability beliefs and interests were asked on a seven-point Likert scale, whereas for University 2 the items were dichotomous. The first limitation of this study lies in the variance of these variables. The low variance might explain why we found that ability beliefs do not add to the prediction of academic success. However, it is also possible that ability beliefs *prior* to enrolment are not predictive of academic success. Van Herpen and colleagues (2017) also found no relation between pre-enrolment academic self-efficacy and obtained number of credits in the first year. Given that the interests-variable is also skewed, but still a significant predictor for non-STEM students and the fact that we found similar patterns in both universities (despite the different response scales), the latter explanation seems more plausible. Hence, we reiterate Van Herpen and colleagues' recommendation for future research to investigate the relation between students' beliefs in their abilities and academic success in a longitudinal setting by measuring ability beliefs both during and after the transition to higher education.

A second limitation of this study is the low explained variance in the outcome measures, but especially in credits. In general, it seems to be difficult to predict obtained number of credits. The variance explained in credits is, just like in our study, usually much lower than the variance explained in (FY)GPA. The number of obtained credits is often used an indicator of re-enrolment (e.g., in the Netherlands, in most programs students are allowed to re-enroll if they obtain 45 out of 60 ECTS) We suggest that future research focuses on predicting retention directly, rather than using number of credits as a proxy for re-enrolment whenever possible.

4.9.2 Implications for Research and Practice

The findings of this study build on previous research that found differences in the predictability of academic success across disciplines (Fonteyne et al., 2017; Willems et al., 2019). These findings imply that it is useful to either control for discipline as a standard practice or conduct more research on specific disciplines to exclude discipline as a possible confounder. By taking such an approach, researchers can extend knowledge of factors that are relevant in the prediction of academic success for students in different disciplines. As a result, administrators could tailor their intake procedures to different groups of students. For example, based on the results in this study, it does not seem relevant to include "interest in the field" in an intake questionnaire for STEM students. Moreover, it appears that the extensiveness of orientation adds to the more known predictors of academic success (e.g., high school performance and conscientiousness). If future research corroborates these findings, administrators could choose to adjust their intake forms and tailor them for specific fields of study.

4.10 Conclusion

This study showed that measuring indicators of academic success prior to enrolment is potentially useful. On top of known predictors like high school performance and personality, indicators of fit such as interest in the program and extensiveness of orientation to the program could also be important in predicting academic success. Moreover, we showed that most of these predictors differed between STEM and non-STEM students. This study shows that differential prediction by discipline might help improve predicting first-year academic success. These results could help university administrators in improving their intake procedures and, as a result, increase their retention rates.

Chapter 5

Pre-university Motivation and First-Year Study Success: Text Mining of Pre-enrolment Questionnaires

Soppe, K. F. B., Bagheri, A., Nadi, S., Klugkist, I. G., Wubbels, T., & Wijngaards-de Meij, L. D. N. V. (2022, submitted). Pre-University Motivation and First Year Study Success: Text Mining of Pre-Enrolment Questionnaires.

ABSTRACT

Student dropout is one of the biggest problems in higher education, because of its negative consequences for both students and universities. Identifying students at risk of dropout prior to enrolment could prevent such negative consequences. Therefore, many selection procedures try to measure student motivation. However, the use of motivation letters has recently been criticized. Staff members are expected to be incapable of classifying students based on their motivation, since the vast amount of text data in the application process exceeds the human ability to process it thoroughly. To alleviate this issue, this study focused on predicting university dropout by using text mining techniques with the aim of exhuming the information contained in students' written motivation. We used machine learning techniques to create new variables (i.e., feature engineering) from the raw text data to enhance the set of characteristics for predicting student dropout. Input for the models consisted of a set of student characteristics, text data (i.e., TFiDF bag-of-words representation), and features extracted from text data (i.e., topics extracted through LDA topic modeling, and cognitive and non-cognitive features from the LIWC text mining tool). A Support Vector Machine (SVM) was used to analyze a sample of 7,060 motivation statements of students enrolling in an open-admission bachelor program at a Dutch university during 2014 and 2015. Results showed that text analysis alone predicted dropout marginally better than a set of student characteristics. However, the combination of text and student characteristics did not improve the prediction of dropout. Suggestions for future research are provided.

Author contributions: KS & AB designed the study. SN preprocessed the data. SN & AB analyzed the data. KS wrote the paper. TW, IK, & LWM provided extensive feedback on all components.

5.1 Introduction

Improving student retention is one of the biggest challenges in higher education. Retaining students results in higher revenue for universities (Zhang et al., 2010) since their funding is often partially based on graduation rates (Jongbloed et al., 2018; Kirk, 2018). For students, finalizing their degree is also of importance, as dropping out of higher education is associated with negative consequences, such as untapped human potential, a low return on their financial investment (Jaeger & Page, 1996; Oreopoulos & Petronijevic, 2013; Psacharopoulos, 1994), or reduced social welfare (Hällsten, 2017). Moreover, low retention rates also impact society since income levels rise with a higher education degree (Jayaraman, 2020). Thus, it is paramount for society to keep dropout in higher education to a minimum.

Identifying students at risk of dropout is complex. Past research has resulted in the identification of many risks (e.g., lacking a sense of belonging, procrastination, adjustment problems) and protective (e.g., high school GPA (HSGPA), conscientiousness, perceived self-efficacy) factors of dropout (e.g., Samuel & Burger, 2020; Tinto, 1987). Based on these factors, a wide array of interventions, such as academic probation or mentoring systems, to prevent dropout has been established in the past, with varying effectiveness (see e.g., Sneyers & de Witte, 2018 for a meta-analysis).

Ideally, students at risk of dropout should be identified prior to enrolment, to minimize negative consequences for both students and universities. In selective admission, it is common practice to try to identify students at risk of dropout based on their application. Staff members of the admissions committee are generally looking for both cognitive (e.g., prior performance) and non-cognitive (e.g., personality and motivation) factors when selecting suitable candidates (Kurysheva et al., 2019). The use of some of these non-cognitive criteria, especially by means of motivation and recommendation letters, for selecting students has been subjected to criticism (Kira Talent, 2018; Posselt, 2016). Self-report measures such as motivation letters are susceptible to faking by the applicant, when being used in a high-stakes context (Niessen et al., 2017). Moreover, filtering out true motivation can be challenging for program staff. They may need to “read between the lines” to form an idea about the factors driving a student to apply for their program. Furthermore, it might be hard to identify students’ motivation solely based on a written statement and characteristics of the reader (e.g., experience), their psychology, and environment can introduce bias into the evaluation of the motivation letters (Bridgeman, 2013). These aspects make humans inconsistent and unreliable evaluators (Zupanc, 2018). Lastly, reading these statements is very time consuming and it is not easy to compare motivation across students.

This study, therefore, focuses on predicting university dropout by using text mining techniques to exhume information contained in students' written motivation. The aim of this study is to investigate whether this novel approach can disclose information present in text, and thereby contribute to detecting students who are potentially at risk of dropout as early as possible. If so, traditional prediction models could be updated, using these techniques, to obtain higher predictive power.

Using machine learning techniques in education is not a new phenomenon. Educational Data Mining (EDM) is defined as "an emerging discipline, concerned with developing methods for exploring the unique and increasingly large-scale data that come from educational settings and using those methods to better understand students, and the settings which they learn in" (International Educational Data Mining Society, n.d.). Many studies in EDM use student registration data, such as high school GPA and demographics in combination with first-year performance to predict various types of academic success (see for examples: Delen ,2010; Hutt et al., 2018; Lauría et al., 2012; Thammasiri et al., 2014).

Thus, almost all EDM research on student dropout prediction makes use of structured data (i.e., quantitative, alpha-numeric data that can directly be used as input in statistical models). There are, however, some studies using Natural Language Processing (NLP) techniques and unstructured data (i.e., qualitative data in no particular format, such as text, audio, or video files) in predicting student completion of Massive Open Online Courses (MOOCs). Most of these studies use sentiment analysis to detect positive or negative phrases, motivation, engagement, etc. in discussion forums or assignments (Jayaraman 2020). For example, in a study on students' opinion towards a course, Wen and colleagues (2014) found, using sentiment analysis, that students who used words related to motivation were more likely to complete the course. Moreover, Crossley and colleagues (2016) used NLP techniques on MOOC forum posts and found that a range of NLP indicators, such as lexical sophistication and writing fluency, were predictive of student completion of the MOOC.

Outside of MOOCs, the authors know of only two studies that used text mining and NLP techniques with unstructured text data to predict dropout. One of these studies used sentiment analysis to predict dropout by analyzing notes written by student advisors (Jayaraman, 2020). By comparing several models, the most accurate model was sought to predict dropout. The poorest performing model, a logistic regression, predicted dropout with 69% accuracy, while the best performing model, a random forest classifier, predicted dropout with 73% accuracy. In the second study known to us, Stone and colleagues (2019) used both human coding and a variety of NLP techniques to detect non-cognitive traits, such as psychological connection (which they used as a proxy for intrinsic motivation), by analyzing students' 150-word open-ended descriptions of their own extracurricular activities or work experiences included in their college applications. Correlations between human coding and model-based coding ranged from medium-small to medium-strong on the respective non-cognitive traits. The correlation between human- and model-based coding for psychological connection was .655, indicating a medium-strong relation. The non-cognitive traits were then used in separate regression models to predict 6-year graduation outcomes. Psychological connection was insignificant in both the human- and the model-coded prediction of 6-year graduation. However, results showed that some other traits had predictive power net of other known predictors. For example, results from both the human- and model-based coding models showed that students portraying a growth mindset were more likely to graduate within six years, when controlling for sociodemographics, secondary school GPA and intelligence.

In this study, having the aim to contribute to early detection of at-risk students, we use NLP-based techniques to analyze short motivation statements of applicants to open-admission bachelor programs in the Netherlands. In doing so, we try to answer the question *whether students at risk of dropout can be identified through text mining, based on their motivation for the program of their initial choice as written in their intake questionnaire prior to enrolment; and whether information extracted from these motivation statements adds predictive power net of student characteristics.*

5.2 Methods

5.2.1 Sample

The dataset is composed of 7,060 motivation statements of students who enrolled in an open-admission bachelor's program during the academic years of 2014 and 2015 at a university in the Netherlands. These motivation statements are part of an online intake questionnaire, which is part of the application process for open-admission programs. Moreover, information on first-year dropout, and student characteristics was obtained from the central student administration.

5.2.2 Measures

In this study both structured data (i.e., a set of student characteristics ranging from prior education to the number of programs a student applied for) and unstructured data (i.e., motivation statements) were used to predict first-year student dropout. Below, we discuss the operationalization of the structured data. In the next sections we will elaborate on the unstructured data and how features were extracted from the raw text data.

Dropout[10]**.** Student dropout is measured using information on whether a student re-enrolled in the second year of the program. Students who paid the tuition fees for the second year are considered to have re-enrolled for their sophomore year. They are treated as the retention class (0). Students who did not pay tuition for their sophomore year are classified as dropouts (1).

Prior education. Students with a variety of educational backgrounds enroll in Dutch universities. Prior education was measured in our sample using a categorical variable with the following levels: "preparatory university" diploma (VWO), to be obtained (1); preparatory university" diploma (VWO), already obtained (2); university of applied sciences propaedeutic diploma (3); and other (4).

High school performance. To measure high school performance (on a scale from 1-10), a mean of the three core subjects in the pre-final year of high school as self-reported in the questionnaire was taken (Dutch, English and Mathematics). If one of the grades was missing, a mean of the remaining two subjects was taken. When grades for only one subject were available, the variable was coded as missing. For students transitioning from a university of applied sciences, grades from their final year were used, for all others grades from the pre-final high school year were used.

Ability beliefs. Students' belief in their own abilities was measured using a single item from the questionnaire, stating: "the program matches my capacities and skills". The item could be answered with (1) yes or (0) no.

Interests. Students' interest in the program was measured using the statement "the program is in line with my interests". The item could be answered with (1) yes or (0) no.

Gender. Information about students' gender was taken from the university registration systems. Male students were coded (1) and female students as (0).

[10]Whether or not students paid for their re-enrolment is one way of operationalizing dropout. Another way is to look at whether students have obtained sufficient credits to continue. In the Netherlands, students receive a Binding Study Advice (BSA) by the end of their freshman year. Generally, a student must obtain 45/60 ECTS to continue in their sophomore year. For this study we ran all models with this classifier as well and results differ only marginally.

Age. Students' date of birth was taken from the university registration systems and then recoded into age at the start of the program by subtracting date of birth from the first day of the academic year for each cohort.

Cohort. The dataset consists of students who enrolled in an open-admission bachelor's program during the academic years of 2014, coded as (1) and 2015, coded as (2).

Discipline. Programs were allocated into three different disciplines: Science, Technology, Engineering & Mathematics (1), Social Sciences (2) and Humanities (3).

Program. To control for differences in dropout across programs, all programs were added as dichotomous variables, i.e., coded as (1) if a student applied for that program and as (0) otherwise.

Previously enrolled. Students who have previously been enrolled in another program at the same university were coded (1) and all others as (0).

Multiple requests. Students who filed more than one admission request (i.e., for more than one program) were coded (1) and all others as (0).

5.2.3 Preprocessing the Motivation Statements

To analyze unstructured text data several steps need to be taken to reduce its high dimensionality. Our raw text data consist of students' answer to the following question on an intake questionnaire: "Why do you want to study [program name] in [city]? (10-25 lines)". This question was followed by the instruction: "When answering the question, for example think about your motivation for the content of the program, your choice for an academic program, and your motivation for a profession or position that this program prepares you for". The first step that needs to be taken is pre-processing the data. In our analysis, pre-processing the motivation statements consists of stop word removal, removing whitespaces and numbers, and converting text into lowercases. This is a step that enhances the performance of the algorithm in later stages. After the pre-processing step, the high dimensionality of text data still prevents it from being used directly in a statistical model and therefore a feature engineering step is required to inspect the text data and acquire multiple feature sets.

5.2.4 *Feature Engineering*

Feature engineering is a process in machine learning that is used to extract analyzable properties from raw data. In machine learning these analyzable properties are known as features; they can be considered independent variables. In this study, three different types of feature engineering were applied to the motivation statements. Initially, a bag-of-words representation was used as a simplified representation of the text data. Thereafter, two additional, and more advanced, feature engineering methods were applied; latent Dirichlet allocation topic modelling, and linguistic inquiry and word count (LIWC) dictionary words. Each of these methods is explained below.

Bag-of-Words. This is a process that converts text data into numbers in a (e.g., document-term) matrix. To create this matrix, we used Term Frequency inverse Document Frequency (TFiDF) which is a bag-of-words method intended to reflect the relative frequency of a term (word) in each document (motivation statement). TFiDF can be calculated by multiplying the number of times a word appears in a document, and the inverse document frequency of the word in the dataset. With TFiDF, the words that are common in every document rank low even though they appear many times. This is because TFiDF is offset by the number of documents that contain the word. The bag-of-words representation is the simplest way to make text analyzable in a statistical model. In the remainder of this paper we will refer to it as just “text” or the method we used: “TFiDF”.

Topic modeling. One way of reducing the high dimensionality of text data, is to represent it as a set of topics across documents. So, instead of looking at the word frequency of single words like TFiDF does, words are clustered into groups that represent underlying concepts. To identify topics in the motivation statements we employed a topic modeling method, Latent Dirichlet allocation (LDA), using a collapsed Gibbs sampling approach (Blei et al., 2003). LDA topic modeling considers each topic as a probability distribution over terms, and each document as a combination of topics. LDA is an unsupervised method, which means that it can automatically extract hidden topics from text without human assistance. This entails that the extracted topics do not always construe a clear meaning. Therefore, it is up to the researchers to identify which number of topics is the best conceptual representation of the text data. Therefore, we ran a set of models with different numbers of topics (i.e., 5, 10, 15, 20, 50). Based on inspection of the representative terms in each topic and to what extent these terms could form a meaningful topic together, two of the authors independently selected the model with 15 topics as the best representation of the data. Therefore, the 15 topics were used as “features extracted from text”.

Linguistic Inquiry and Word Count (LIWC). LIWC is a text analysis tool that reduces the dimensionality of the data by mapping words onto predetermined categories, using psychometrically validated dictionaries. LIWC can expose certain psychological characteristics of the writer (Tausczik & Pennebaker, 2012). The main categories provided by LIWC are general information (e.g., word count), linguistic dimensions (categorized in verbs and function words, such as pronouns), psychological processes (containing the main categories social processes, affective processes, cognitive processes, perceptual processes, biological processes, and relativity), personal concerns, and spoken language (Pennebaker et al., 2007). Each of these categories has some features. For example, the category "social words" contains the features "family", "friends", and "humans", and the category "relativity" is divided into the features "motion", "space" and "time". We use LIWC2007 to extract these features from the motivation statements and to use as input for the models. In this version of LIWC, its dictionaries have been translated into several languages, including Dutch (Boot et al., 2017). Together with the 15 topics, we refer to the LIWC sub-categories as "features extracted from text" in the remainder of this paper.

5.2.5 *Training the Algorithm*

Several machine learning techniques can be used to analyze the data. Support Vector Machine (SVM) was chosen since it generally performs well in text classification problems. It is beyond the scope of this paper to compare the performance of the SVM to other techniques, like K-nearest neighbors or naïve Bayes (see for example, Aggarwal et al., 2012; Kowsari et al., 2019, for text mining studies on comparing the performance of different techniques). First, the data were split into training (75%) and test sets (25%). The training set was used to train the algorithm by providing it with both the input (i.e., student characteristics, text, and text features) and the output (i.e., whether a student dropped out or not). K-fold cross validation with k = 5 was used to evaluate the performance of the training model (Refaeilzadeh et al., 2009). Cross validation is a resampling process in which the training data is split in k-different portions to test and train a model on different iterations. If the different iterations return different levels of model accuracy, this can indicate potential problems regarding overfitting or selection bias. The full set of training data can then be inspected further to ensure better performance of the algorithm when eventually applied to the test data. For none of the estimated models in our study cross validation indicated a need for further inspection of the training data.

5.2.6 *Analysis*

We analyzed our data using six separate models, exploring the most accurate combination of measures to predict dropout. First, we started with a model using only the structured data, our set of student characteristics, as input for the model. This provided us with a baseline of how well we would be able to predict dropout if we would not have any text data. Second, we estimated a model with only the text using TFiDF as input to compare the algorithms' performance to that of the first model. Third, we added the features that we extracted from the text through LDA topic modeling and the LIWC dictionary to the text-only model to assess the added value of more advanced text mining techniques on top of the simple bag-of-words representation. Lastly, to answer the question whether information extracted from text can add to the prediction of dropout net of structured data, we examined the performance of the SVM with different combined feature sets. Table 5.1 provides an overview of the input per model.

Table 5.1 Input for the estimated models to predict dropout.

Model	Input
1	Student characteristics
2	Text (TFiDF)
3	Text (TFiDF) + features extracted from text, using LDA & LIWC
4	Student characteristics + text (TFiDF)
5	Student characteristics + features extracted from text, using LDA & LIWC
6	Student characteristics + text (TFiDF) + features extracted from text, using LDA & LIWC

Student retention data are generally imbalanced, since the number of dropouts is much smaller than the number of students that continue the program. This imbalance can be problematic, as the standard classification algorithms have a bias towards the majority class, giving misleadingly promising results (Dalipi et al., 2018). The technique we used to correct this imbalance combined oversampling the minority class and undersampling the majority class in such a way that the model automatically assigns the class weights inversely proportional to their respective frequencies.

Performance of Support Vector Machines is generally assessed by accuracy, precision, recall, and f1-scores. Accuracy is an unreliable measure for imbalanced data and therefore we do not use it. Moreover, because of the imbalance, weighted output for precision, recall, and f1-score are reported to assess the performance of the algorithm. Precision denotes the true positives divided by the true positives + false positives. Recall is defined as the true positives divided by the true positives + false negatives. The f1-score is the weighted average of precision and recall. To get a better sense of these performance measures for our specific context, Figure 5.1 shows precision and recall for the dropout class.

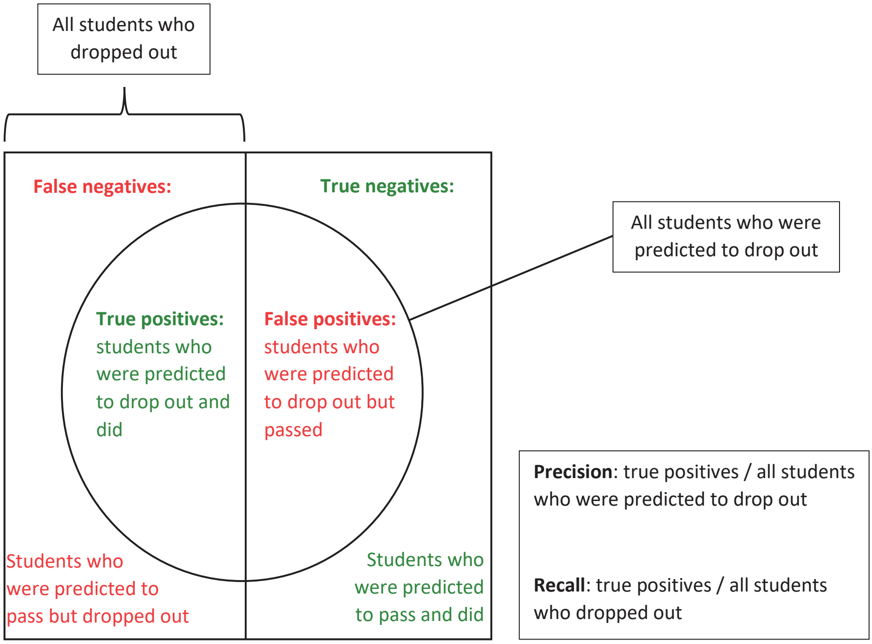

Figure 5.1 Precision and recall for the dropout class.

5.3 Results

Results presented in this section are based on 1765 motivation statements (25% of the total dataset), together forming the test data. Of the 1765 statements, 1312 belonged to the retention class and 453 to the dropout class. Table 5.2 provides a summary of the results.

Table 5.2 Output of the estimated models, for the total sample (T) and split by the retention (R) and dropout (D) class.

	Precision			Recall			f1-score		
Model	T	R	D	T	R	D	T	R	D
1	.71	.84	.31	.57	.55	.65	.60	.66	.42
2	.67	.79	.28	.63	.71	.37	.65	.75	.32
3	.67	.79	.29	.64	.72	.37	.65	.75	.32
4	.69	.81	.32	.65	.71	.44	.67	.76	.37
5	.70	.82	.31	.60	.61	.56	.63	.70	.40
6	.68	.80	.31	.64	.71	.42	.66	.75	.35

5.3.1 *Model 1: Student Characteristics*

First, a model with only student characteristics was run as a baseline model. This model contains variables that are known to be predictive of student dropout, including high school performance. The weighted precision score of Model 1 was .71, the weighted recall score was .57. This resulted in a weighted f1-score of .60 for this model. The prediction of precision for the majority class (retention) was much better than the performance for the minority class (dropout). With a score of .65, recall was higher for the dropout class than for the retention class (i.e., .55). This is notable, given the fact that algorithms generally perform better for the majority class.

5.3.2 *Model 2: Only Text*

To identify what the analysis of text data through NLP-based techniques can add to the prediction of dropout, first a model with only text was analyzed, using TFiDF to identify the importance of words in the corpus (Model 2). The model had a weighted precision score of .67, a weighted recall score of .63, and a weighted f1-score of .65. When comparing the performance of the algorithm for this model several things stand out. First, precision of this model is worse than in Model 1, meaning that in the model with the student characteristics, of all the selected students there were proportionally more true positives. In Model 2 recall is better than in Model 1, meaning that of all the relevant students there were proportionally more true positives in the model with only text. Second, recall scores for this model are less balanced across the classes than in Model 1. Table 5.2 shows a much higher recall score for the retention class (i.e., .71) than for the dropout class (.37).

5.3.3 *Model 3: Text and Text Features*

Model 3 investigates the predictive power of all information we extracted from text. To that end, we added features extracted from text to the text data (LIWC dictionary words and topics extracted through LDA topic modelling), to investigate whether this combination could outperform Model 1. The LIWC features already consist of predetermined categories, using psychometrically validated dictionaries. For the LDA topic modelling the features first must be identified before they can be used in the analyses. For the 15 topics that were identified in the motivation statements, the top 10 terms of each of these topics are listed in Table 5.3. Upon consensus among the authors, the 15 topics were given a theoretical label if possible. Some of the topics did not construe one clear underlying concept and were therefore left unlabeled.

Table 5.3 Top ten terms per topic.

Topic	Label	Top words
1	program interest	study, Utrecht, program, very, finds, fun, seems, good, very, rather
2	general interest	university, highly, knowledge, Utrecht, study, interest, offers, like, choice, developing
3	previously enrolled	year, study, go, wanted, came, found, rather, choice, studying, knew
4	applied sciences	program, university of applied sciences, Utrecht, scientific education, university, less, aspects, academic, difference, choose
5	culture minded	subjects, different, language, interests, culture, year, broad, cultures, choosing, liberal arts
6	sense of belonging	day, open, trial studying days, visited, found, open days, during, ambiance, spoke, immediately
7	societal minded	social, spatial planning, study, sciences, general, geography, studies, different, expect, hope
8	pedagogical minded	children, study, pedagogical sciences, chosen, helping, doubts, good, characteristics, later, finished
9	computer minded	programming, games, computers, artificial intelligence, game technology, suitable, technology, game, computer, logic
10	artistic students	media, art, theater, culture, film, television, films, chose, theater film, Amsterdam
11	-	music, nature, hopefully, astronomy, most important, therein, teaching school, ever, madam, dear sir
12	-	person, maximum, character, getting acquainted, function, fascinated, legal system, honest, nature (kind), wonderful
13	location	location, widening, exist, per, passed, stadium, analysis, classes, acquaintances, about
14	politics minded	political, strike, stone (figure of speech), technological, horizon, sustainable, advance, curriculum
15	-	help, automatically, job opportunity, sociological, public, mono disciplinary, suits

Figure 5.2 shows the feature importance of the different topics. The topics 13 (i.e., location), 7 (i.e., societal minded), and 11 (unlabeled) were the strongest positive predictors of dropout. Students using a lot of words related to these aspects, are more likely to drop out. The strongest negative predictors of dropout were the topics 2 (i.e., program interest), 12 (unlabeled), and 1 (i.e., general interest). Students using many of such words, are less likely to drop out.

Subsequently, all 15 topics were fed to the SVM, together with text data of Model 2 and features from the LIWC package for text analysis. The results in Table 5.2 show that Model 3 is almost identical to Model 2. In other words, it appears that, in this study, the features extracted through LDA topic modelling and the LIWC package do not add to the prediction of dropout in comparison to text alone.

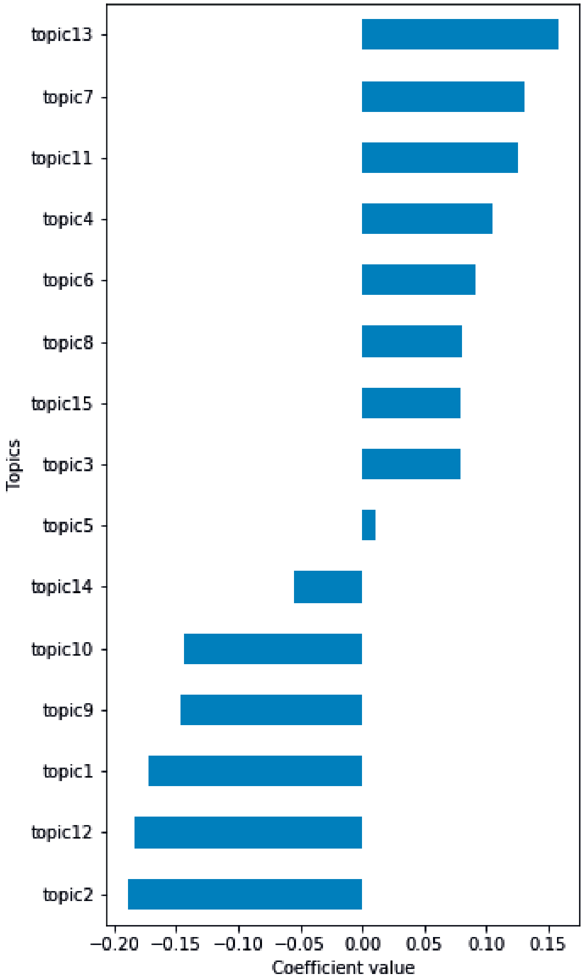

Figure 5.2 Feature importance of the LDA topics.

5.3.4 *Model 4: Student characteristics and text*

To identify whether written motivation adds to the prediction of first-year dropout net of student characteristics, Model 1 was combined with different elements of text. In Model 4 the input of Model 1 (student characteristics) and Model 2 (text only) was combined. The weighted recall (i.e., .65) and weighted f1-score (i.e., .67) in this model are the highest of all the estimated models. The weighted precision score (i.e., 69) of this Model holds a middle position regarding algorithm performance between Model 2 and 3 on the one hand, and Model 1 on the other hand.

5.3.5 Model 5: Student Characteristics and Text Features

In this fifth model, student characteristics were combined with features extracted from the text, rather than the text itself. Even though the features extracted from text did not add to the predictive power of dropout net of text alone (Model 3), the combination of student characteristics and features extracted from text might improve the prediction of dropout. With a weighted precision score of .70, a weighted recall score of .60, and a weighted f1-score of .63, this algorithm performed worse than in Model 4.

Model 5 was also used to inspect the importance of all features together that were used as input for the SVM (i.e., student characteristics, LIWC words, and LDA topics). We performed this step on this model, rather than Model 6 (see below), because the vast number of features of the TFiDF method (i.e., all the individual words in the texts) does not allow it to be captured in a figure. The importance of the 25 most important features for dropout is shown in Figure 5.3. Some of the most important features are discussed. The strongest positive effect is word count (WC), indicating that the more words students used the higher their probability of dropout. Second, the use of personal pronouns (ppron) is a similarly strong predictor of dropout in this model. The more personal pronouns a student uses in their text, the higher the probability of dropout. Frequent use of the first person singular (i), however, is negatively related to dropout. Looking at Figure 5.3, it indeed seems to matter which pronouns are being used. For example, the more the second person (you) is used, the lower the probability of dropout, whereas the relatively frequent use of impersonal pronouns (ipron) is associated with a higher probability of dropout. It is noteworthy that, in this model, high school performance was again the strongest negative predictor of dropout. Lastly, relatively important features in this list are age and article. Age was a relatively weak predictor in the model with only student characteristics. In this model, however, it is the third most important predictor of dropout, with older students having a higher probability to drop out. The relatively frequent use of articles (article), on the other hand, is associated with a lower probability to drop out. Among the top 25 most important features for the prediction of dropout there are several other features that were obtained through the LIWC dictionary. Interestingly though, none of the topics from the LDA topic modelling is amongst the 25 most important predictors of dropout.

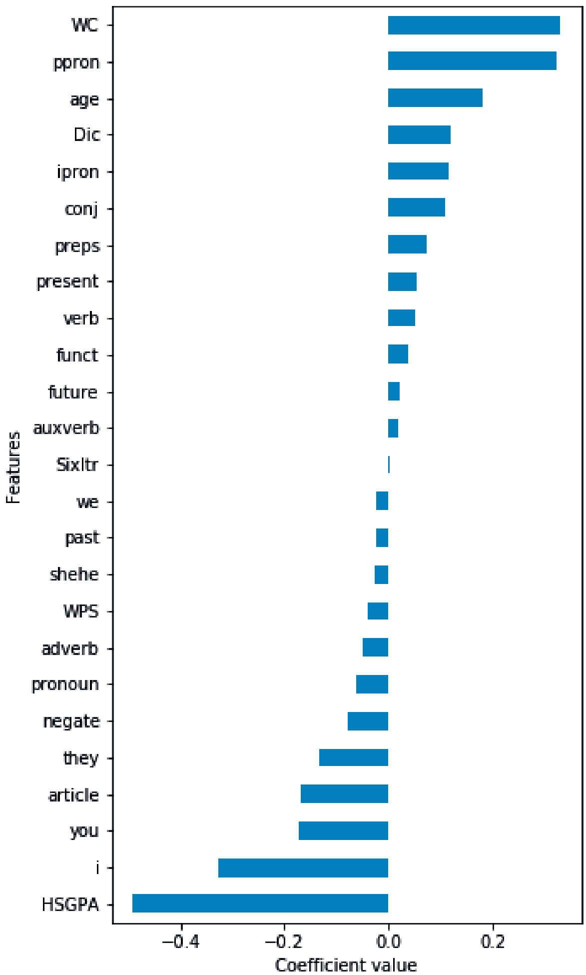

Figure 5.3 Feature importance[11] of the top 25 features from Model 5.

5.3.6 Model 6: Student Characteristics, Text, and Text Features

Lastly, a model was estimated that included the student characteristics, as well as the text itself and the features extracted from the text. The algorithm performance of this model was almost the same as the performance for Model 4. The weighted precision was .68, weighted recall was .64, and the weighted f1-score of this model

[11]Meaning of the features from top to down: word count (WC); personal pronouns (ppron); age; dictionary words (Dic); impersonal pronous (ipron); conjunctions (conj); prepositions (preps); present tense (present); common verbs (verb); total function words (funct); future tense (future); auxiliary verbs (auxverb); words with more than 6 letters (sixltr); first person plural (we); past tense (past); third person singular (shehe); words per sentence (WPS); adverbs (adverb); total pronouns (pronoun); negations (negate); third person plural (they); articles (article); second person (you); first person singular (i); HSGPA (high school performance).

was .66. When comparing this model to Model 5, it strengthens the conclusion that features extracted from text are, in our sample, of limited additional value in the prediction of dropout.

5.3.7 *Comparing Most Frequently Used Terms of Correctly and Incorrectly Classified Dropouts*

To get a better understanding of the algorithm performance, top terms used by students who were correctly identified as dropouts, were compared to top terms used by students who were incorrectly classified as dropouts. A high overlap in the commonly used terms, would indicate that there is not enough discrepancy in the written motivation between the two groups for the SVM to detect any differences.

When we inspected the 100 most used terms for both groups, overlap was indeed identified. Roughly a quarter of the top terms was used by students in both categories. Most of these words were names of programs (e.g., *Biology, Law, Sociology*), or derivatives thereof (e.g., *game(s)* for Game Technology or *art* for Art History). The other overlapping words are generic, such as *program* or *name*, or apply to a specific field (i.e., *people* and *children* for behavioral sciences and *music* and *film* for arts programs). Given that most of the overlapping words refer to names of programs or derivatives thereof, the prediction of dropout may improve if these words can be excluded from the text input. Because of the too small sample size per program in our data we were not able to do this.

5.4 Discussion

In this study we attempted to answer the question whether students at risk of dropout can be identified based on their motivation for the program of their initial choice as written in their intake questionnaire prior to enrolment by using NLP-based techniques. Moreover, we asked the question whether information extracted from these motivation statements adds predictive power net of student characteristics. The results showed that the answer to this question is twofold. When text was used in addition to student characteristics, it hardly added to the prediction of dropout. However, when only text data were used, the algorithm performed very similar to the one in the model with only the student characteristics.

Predicting dropout accurately is not easy, especially not based on student information that is available prior to enrolment in higher education. Since the model with only text showed very similar results to the model with only student characteristics, it appears that student dropout can be predicted with a short motivation statement analyzed with data mining techniques at least as good as with a set of known predictors like high school performance. Moreover, these two types of predictors seem to complement each other, as precision was higher in the model with student characteristics (Model 1) and recall was higher in the model with only text (Model 2). Therefore, analyzing text data with text mining techniques seems promising. Our aim was to exhume hidden information from text data and investigate whether this information could be used to predict students at risk of dropout. Unstructured data, like text, are very time consuming and complex to analyze for humans. However, if highly predictive text mining algorithms can be developed to analyze these data, that could potentially be useful to identify students at risk of dropout before the start of the program without needing an extensive set of student characteristics. Such students could then immediately be offered support to mitigate the risk.

The fact that combining the text and student characteristics, like high school performance, does not (substantially) improve the model prediction in this study, might indicate that they measure the same underlying concepts. It is possible that the way the question about motivation for the program was asked or explained, probes students to put into words the information they already filled out earlier in the questionnaire by answering the other survey questions. Future research could try to verify this hypothesis by studying motivation statements using less directive questions. Another possible way might be to ask more open-ended questions about specific components of the program choice (e.g., why this program; why this university; what in this program makes it attractive; etc.) to obtain more unstructured (i.e., text) data covering a variety of underlying concepts to analyze and relate to academic outcomes, using machine learning techniques.

A limitation of this study lies in the properties of our sample. First, the dataset is imbalanced, because there are proportionally few dropouts. This is generally the case with student retention data, and therefore, cannot be solved. Oversampling and undersampling techniques were used and weighted scores were reported to deal with this limitation. Second, the motivation statements are generally short (i.e., students were requested to write 10-25 lines, resulting in texts that are roughly 250 words long) and the sample consists of applicants to all open-admission bachelor programs. Both the length of the texts and the heterogeneous sample may have an influence on the ability of the algorithm to construct an accurate prediction model. Algorithms learn by providing them with more and more data. Despite our relatively large number of motivation statements, the relatively short and pluriform texts that were used could have affected the performance of the algorithm for the text models. Future research may investigate whether a more uniform sample (e.g., of one faculty or one program) would result in a better performance of the text mining approach.

Another direction for future research is to apply deep learning-based NLP methods with the use of transfer learning (i.e., improving learning in a new task through the transfer of knowledge acquired in a previous task) on a bigger dataset. This could improve representation of text data using the distributional hypothesis, which poses that the more semantically similar two words are, the more distributionally similar they will be, and thus the more they tend to occur in similar linguistic contexts (Sahlgren, 2008). For example, the algorithm can use the fact that the words *city* and *location* are distributionally more like one another than they are to the word *scientific* in a multidimensional context, to predict that *city* and *location* are also semantically more like one another than to the word *scientific*. However, these techniques require more data. Nevertheless, this direction is worth researching as it could help in capturing the distinctive writing style of a student. This information could, in turn, contribute to the early identification of dropouts.

When developing prediction models with the aim to use them for the early identification of dropouts, one should especially focus on improving the precision scores for the dropout class. If text mining methods were to become reliable enough to be used in advising students on their program choice, it would be better to incorrectly classify a student as successful, than to incorrectly classify a student as a future dropout. Some students who were then incorrectly advised to start the program might actually be successful if the positive advice has an influence on their feelings of self-efficacy and/or motivation. Regardless of whether an advice can have such an effect, it is better to have a predictive instrument that returns very few false positives (students unjustly classified as dropout) than an instrument that returns very few false negatives (students who were unjustly classified as successful), because staff are generally in favor of giving every student an opportunity to study in higher education. Therefore, if choices must be made in developing these models, prioritizing a high precision score for the dropout category is most favorable when you are wanting to give every student a chance.

There is a famous quote by the economist Ronald Coase stating: *"if you torture the data long enough, it will confess to anything"*. Coase meant this as a warning to fellow researchers, not to engage in scientific misconduct (e.g., p-hacking). Although not for the purposes of making it confess to anything specific, torturing the data is exactly what we did in this study. We approached the motivation data from different angles, to predict first-year dropout of students applying for a certain undergraduate program. By comparing and combining different methods, we found that applying machine learning techniques on student motivation is a potentially promising new way of approaching student dropout. We believe it is worthwhile to explore this line of research further to find better ways to extract information from motivation statements. When doing so, focus could also be placed on comparing human-coded and algorithm-coded motivation statements to get a better sense of how accurate these methods are in predicting dropout and which of them is better.

Chapter 6

Discussion

6.1 Introduction

The aim of the research presented in this dissertation was to evaluate the effectiveness of various types of matching procedures in Dutch higher education. These procedures are implemented to serve as a final check on students' program choice and are aimed at making students enroll in a fitting program. The research questions in this dissertation were approached from a perspective of person-environment fit, which is defined as the compatibility between individual and environmental characteristics (Kristof-Brown & Guay, 2011). Within the educational context, person-environment fit builds on the assumption that students with certain characteristics are more likely to choose certain programs (Astin, 1993) and that congruence between student and program is paramount to academic success (Feldman et al., 1999). Throughout this dissertation we argued that students who can test their fit with the program of their choice prior to enrolment will make a better choice. Using input from several motivational theories, we identified three concepts that we deemed important in the context of a higher education program choice. The first concept that we identified is ability beliefs, which was defined as "one's beliefs in their abilities to perform a certain task". Second, we defined the concept interests as "the extent to which a person values certain topics over others". And third, the concept sense of belonging was defined as "a sense of connectedness with fellow students, staff members and one's physical surroundings".

This chapter provides a summary of the findings presented in previous chapters, followed by a discussion of the scientific contributions of the research presented in this dissertation. Then, we reflect on the practical implications of the results for matching procedures in the Netherlands. Thereafter, we address the limitations of the research in this dissertation and provide suggestions for future research. Lastly, we present some concluding remarks.

6.2 Summary of the Main Findings

In this dissertation we researched the effectiveness of various types of matching procedures as experienced by students (Chapter 2) and in relation to enrolment (Chapter 3) as well as first-year academic success (Chapter 4 and Chapter 5).

In Chapter 2 we researched student perceptions of different elements of Dutch matching procedures. Matching procedures serve as a final check on students' program choice. We interviewed 61 prospective students of several programs at four different universities about their perceptions of the elements of matching (i.e., questionnaire; activity in the form of a personal interview, online course, or matching day; feedback or advice) in relation to their program choice. The results showed that students perceive those matching procedures as potentially able to contribute to providing insight into the program, as well as to confirming their program choice. However, the extent to which matching can contribute to these aspects differs between elements of the matching procedures, and between groups of students. Interestingly, the two elements that are part of the matching procedure at (almost) every university program, i.e., the questionnaire and the advice, were deemed least useful to prospective students in their program choice. Matching activities (i.e., personal interview, online course, or matching day) were considered more useful in making a final program choice than the questionnaire and the advice, but there were differences in students' perception of usefulness between the three types of activities. Overall, the more aspects of person-environment fit a student could test, the more useful the activity was regarded.

In Chapter 3 we studied the relation between types of matching procedures and differences in enrolment rates of thirteen programs at four Dutch universities. We argued that the lower enrolment rates, the more students had changed their mind after participating in a matching procedure. Findings of this research showed that more intensive matching procedures, in which more aspects of person-environment fit can be tested, are associated with lower enrolment rates. We showed that the implementation of matching procedures in general, and the intensity of these procedures in particular, are a likely explanation for the lower enrolment rates, because enrolment rates have dropped since the implementation of matching for programs with intensive procedures, but not for programs that only offer follow-up activities for students at risk.

In Chapter 4, we investigated the relation between indicators of fit between student and program, as measured prior to enrolment in the matching questionnaires, and first-year academic success. This was done for all open-admissions programs of two universities across three cohorts. Moreover, we explored whether the prediction of academic success differs across disciplines. For this study, we first estimated a Structural Equation Model (SEM) on the data of one university and then replicated the findings using data from a second university. We showed that high school performance, conscientiousness, pre-university program interest, extent of orientation to the program, age, and gender are predictive of first-year grade point average (GPA) and earned credits. Although effect sizes are small, we find almost identical patterns across the two universities in this study. This strengthens our believe that first-year academic success can be predicted using indicators of fit, measured prior to enrolment. A second important finding is that most of the fit indicators in our study differ in strength and/or direction between Science, Technology, Engineering, and Mathematics (STEM) and non-STEM programs. These findings provide indications that non-cognitive indicators (e.g., conscientiousness and program interest) are stronger predictors of first-year academic success for non-STEM students than for STEM students. For STEM students, high school performance is the strongest predictor of first-year academic success.

In Chapter 5 we used motivation statements of matching questionnaires from the applicants to all open-admission programs in one university across two cohorts to investigate whether text-mining techniques can be used to predict first-year dropout. Results showed that the answer to that question is twofold. Motivation statements in the intake questionnaires predict first-year dropout equally well as a set of student characteristics that includes the fit indicators from Chapter 4. However, the combination of text and student characteristics did not improve the prediction of dropout. Thus, on the one hand the use of text-mining techniques for the prediction of dropout seems promising. On the other hand, the fact that combining the text and numeric data did not improve the prediction of dropout in this study, might indicate that they measure the same underlying concepts.

6.3 Effectiveness of Matching Procedures

As discussed in Chapter 1, in this dissertation matching procedures are deemed effective if the procedure 1) is considered useful by students in their final program choice, 2) makes students who are deemed at-risk of dropout reconsider finalizing their enrolment, and 3) is positively associated with first-year academic success. The effectiveness of the Dutch matching procedures will be discussed below according to each of these criteria.

6.3.1 *Student Perceptions of Usefulness*

In Chapter 2 we interviewed prospective students on the role of matching procedures in their program choice. From these interviews we learned that prospective students find the questionnaire and the advice the least useful elements of the procedures. Although some students said that the questionnaire sparks reflection on reasons to choose the program, about half of the respondents in our study indicated that the questionnaires have no added value. A reason for this lack of value could be that it is easy to give socially desirable answers to ensure receiving a positive advice. Moreover, prospective students sometimes expressed that they felt that the questionnaires were mainly for program staff to assess whether a student would fit with the program rather than for the students to test the program-fit themselves. Advice as a result of the matching procedures was also deemed not very useful for the final program choice. Many students indicated they were already certain of their program choice, so the advice was at most a confirmation of their own perceptions. Negative advice was often ignored, for example because students felt that they would know better what would fit them than program staff. Other students indicated that they felt the follow-up activity was more indicative of fit than the questionnaire or they admitted not having taken the questionnaire seriously.

In contrast, matching activities were deemed more useful than the questionnaire and advice in making a final program choice and overall we can say that the more aspects of person-environment fit a student could test, the more useful the activity was deemed. Two interesting deviations from this overall pattern were found. First, of students participating in a matching day, on average, STEM students indicated more often that they had difficulty to assess whether they felt socially comfortable than non-STEM students. This might indicate that sense of belonging plays a different role in the program choice of STEM students in comparison to non-STEM students. Second, participants in two different online matching procedures expressed the usefulness of such a matching procedure for testing their ability beliefs differently. The underlying factor here was their perception of representativeness of the online matching course. Students in one program experienced the online course as too easy and “probably not representative of the courses in the program”. Hence, they found it hard to assess whether they would have the ability to pass the courses in the program. Students in the other program had no such difficulties with testing their ability beliefs. Thus, it should be noted that experience of representativeness for the program is very important for how useful matching procedures are perceived by prospective students in their final program choice.

6.3.2 Making Students at Risk of Dropout Reconsider Their Program Choice

Matching procedures can make prospective students change their mind, provided that enough elements of person-environment fit can be tested. Chapter 2 sheds light on reasons why students may or may not choose to enroll after participating in the matching procedures. The main reason students give for not changing their mind is that the choice had already been made prior to their participation in the matching procedure. Some other reasons are that they think it is too late to switch to another program, they need this specific program for a future job they want, or that they think they can estimate their fit better than the program staff. In our sample there were no students who radically changed their mind from one program to another as a result of the matching procedure. We did speak to quite some students who decided not to pursue their back-up anymore, because they felt a good fit on sense of belonging or interests with the program of their first choice (or a better feeling of fit in one program than the other). Others posed that they might have changed their mind if they would have felt a mismatch in their ability beliefs (e.g., STEM or English harder than expected), interests (e.g., different content than expected), or sense of belonging (e.g., lacking a social connection). Furthermore, in Chapter 2 we found preliminary indications that sense of belonging may play a different role in the program choice of STEM students in comparison to non-STEM students. In Chapter 3 we found additional support for this assumption. When comparing enrolment rates of different types of matching procedures, regardless of institution, we found that the lowest enrolment rates were found in programs that employed a matching day. However, for the STEM program in our sample the finding was reversed; enrolment rates were lowest in programs with an online matching course. In combination with the findings of Chapter 2, this makes us believe that online matching procedures may potentially be more useful for STEM students than matching days and vice versa for non-STEM students. In conclusion, it seems probable that matching procedures can make students re-consider their program choice. However, in general it appears that this happens more in more intensive matching procedures and only if prospective students feel that the matching procedure provided them with a realistic view of the program.

6.3.3 Improving First-Year Academic Success

The relation between types of matching procedures and academic success was not studied in this dissertation (see limitations and future research), but other sources may shed some light on this relation. The most important indicator that there may be an association between types of matching procedures and academic success stems from trends in retention rates (i.e., re-enrolment in the second year of the program) at the institutional level. When inspecting these retention rates, two universities stand out. At the University of Amsterdam and Utrecht University, a clear and persistent increase in retention can be identified in the year the matching procedures were implemented (VSNU, n.d.(a)). These are the only two universities that implemented intensive matching procedures (i.e., both consisting of online intake questionnaires, matching days, homework, and a test) in a uniform approach for the whole institution (VSNU, 2017). Apart from the top-down implemented, uniform procedure, the University of Amsterdam and Utrecht University stand out for making participation in the matching procedures mandatory for every student. Students who do not participate in the matching procedures are not allowed to enroll. Eindhoven University of Technology should also be mentioned in this context (Warps et al., 2017). Since 2009 Eindhoven University of Technology has been conducting research on retention among their bachelor students. Their intake questionnaire is the result of this research. This questionnaire is very extensive and was already in use before the implementation of matching. That might explain why a significant increase in retention occurred in 2012, but not upon the official implementation of matching in 2014 (VSNU, n.d. (a)). There are several other indicators that there may be an association between types of matching procedures and first-year academic success. For example, several universities (e.g., University of Amsterdam) consciously avoided incorporating personal interviews in their matching procedures, because earlier research within their institutions had shown that there was no relationship between these interviews and academic success (Warps et al., 2017). Moreover, Eindhoven University of Technology found no effects of matching interviews on academic success but did experience positive effects on forming realistic expectations and sense of belonging among their students (Warps et al., 2017). None of these indications of a possible relation between types of matching procedures and first-year academic success can currently be supported sufficiently by empirical research, so future research may try to focus on answering this question.

6.4 Scientific Contributions

The studies conducted in this dissertation contribute to the knowledge of how testing fit prior to enrolment can guide students in finding a suitable program and improving their first-year academic success. One important contribution lies in the development of a conceptual model of person-environment fit for the educational context. In Chapter 2 we integrated and built on several theories, of which two motivational theories, Expectancy-Value Theory (EVT; Eccles & Wigfield, 2002) and Self-Determination Theory (SDT; Deci & Ryan, 1985), as well as Tinto's Student Integration Model (1977) are the most important. We identify three concepts that we assume to be important in the context of a higher education program choice: ability beliefs, interests, and sense of belonging. Throughout this dissertation we have shown that the process of choosing a university program can be understood through the lens of testing fit on ability beliefs, interests, and sense of belonging. The application of the person-environment fit perspective to the educational context is an innovative way of approaching the transition into higher education that is not bound to the Dutch educational context. Regardless of the educational system, testing fit prior to enrolment might be a valuable addition to the transition process. In fact, the implementation of a system in which prospective students can test their fit with a future study program might be especially valuable in countries where tuition fees are high, like the UK and the USA.

Second, results in this dissertation indicate that first-year academic success can be predicted using pre-university indicators of fit. Measuring these indicators prior to enrolment, rather than during the first academic year, can contribute to the early detection of students at-risk of dropout. Most of these predictors will also be relevant for other educational contexts. Thus, with the specific characteristics of the educational system in mind, matching procedures can also be developed for other countries. It can be worthwhile to invest in such program choice interventions to improve student-program fit and reduce dropout. However, the design of these interventions is crucial for their effectiveness. The most important outcome of this dissertation in that regard is that interventions should provide prospective students with a realistic view of what the program entails. Moreover, the prediction of first-year academic success may improve if specific indicators are used for specific groups of students. We showed in this dissertation that for STEM students high school performance is the strongest predictor of first-year academic success, while for non-STEM students non-cognitive characteristics (e.g., conscientiousness) are also strong predictors. Another way to improve the prediction of student dropout may lie in applying text-mining techniques in the educational context to extract information contained in written text. We showed that a short, written motivation statement is equally predictive of first-year dropout as a set student characteristics that is known to be predictive of dropout, including high school GPA. Although combining the student characteristics and the motivation statements did not improve model performance in our sample, it appears that they do complement each other based on which criteria are used for the assessment of model performance. Thus, we believe that using text mining and other machine learning techniques may advance the understanding of student dropout.

6.5 Practical Implications

The results presented in this dissertation contribute to our knowledge of what works in improving students' program choice. These results have practical implications for higher education procedures in practice. Below, three suggestions are given for policymakers, university boards, faculty, and researchers involved in (the transition into) higher education in the Netherlands and countries with similar educational systems. The suggestions are related to the three main elements of the current matching procedures in Dutch research universities: intake questionnaires, matching activities, and advice.

First, the intake questionnaires should contain only evidence-based predictors of academic success. In Chapter 4 we have shown that high school performance, conscientiousness, pre-university program interest, extent of orientation to the program, age, and gender are predictive of first-year GPA and earned credits. Moreover, in Chapter 5 we have shown that a written motivation statement is equally predictive of first-year dropout as the combination of the abovementioned characteristics, but these two ways of prediction of drop out seem to complement one another. Therefore, these characteristics and a written motivation combined could form a good start for intake questionnaires. Furthermore, intake questionnaires may be adjusted depending on the field of study or redesigned in such a way that they contain a general and a program-specific part. Based on the findings in our study in Chapter 4, it appears that non-cognitive predictors are important in the prediction of first-year academic success of non-STEM students, while high-school GPA is the most important predictor for STEM students. Other research shows that there may be additional pre-enrolment indicators that could be included in these questionnaires (see for example Van Herpen et al., 2019 who showed that pre-university effort is a predictor of academic success). However, it is advisable to keep the questionnaires short, because prospective students don't find them useful in their program choice. Therefore, they are mainly a means for program staff to assess students' risk of dropout. If program staff uses student answers to these questionnaires in their matching advice and prospective students act accordingly, the questionnaires can contribute to first-year academic success in Dutch higher education. However, for students to act on the matching advice, the system of advising students may also need to undergo some changes (see below).

Second, matching activities should be designed in such a way that they allow for testing person-environment fit. This has several concrete implications. First, a matching procedure should consist of more than only an intake questionnaire and advice. Matching is first and foremost designed to help students in their program choice and both the intake questionnaire and advice are not considered useful in this process by prospective students. The more aspects of person-environment fit can be tested in a matching procedure, the more useful prospective students find them. In our research, these were online courses and, especially, matching days. Given the importance of sense of belonging in (the transition into) higher education, programs offering online matching procedures should consider whether they can incorporate any form of contact between prospective students within their procedures, for example in the form of establishing some sort of chatroom or a Q&A session with current or prospective students. A third matching activity that we studied in this dissertation concerned personal interviews. Although prospective students experience interviews as a relatively useful resource in their program choice, they are time-intensive and, therefore, expensive. Moreover, previous research has shown that interviews are bad tools for predicting first-year GPA (e.g., Dana et al., 2013; Reumer & Van der Wende, 2010). Given their lack of predictive power, time and money dedicated to conducting personal interviews with applicants can probably be better spend differently.

Third, the system of advising students may be reconsidered. Although there is a clear association between a negative advice and student dropout, many students ignore the negative advice (Warps et al., 2017). This is especially the case, if prospective students have the feeling that the matching is not representative of the program (Soppe et al., 2019, Chapter 2). One student in our study in Chapter 2 formulated it as follows: *"I do actually think that I can handle this study. When I see…Yeah, I think, say, I personally didn't think I had done very badly in the matching class either. So, I had more than half the points anyway, so I also didn't understand very much why they thought I needed a negative recommendation. But it didn't affect me very much"*. Overall, program staff hardly gives negative advice and an advice that expresses doubt is uncommon as well. Leenheer (2022) formulates this as follows: *"The matching officer does not aim to be very strict, but rather to filter out those applicants for whom serious doubts exist"*. Program staff wants to send a signal with this advice, but precisely students who receive a non-positive advice find it often unjust and not useful (Warps et al., 2017). That it is precisely students with a non-positive advice who ignore it, can be explained by the fact that the way in which feedback is perceived, strongly determines whether it is acted upon (Neimeijer, 2020). For example, Van Gurp and Van den Hurk (2014) found that negatively formulated feedback could influence self-confidence, resulting in feedback being perceived as not useful. Moreover, Vulperhorst and colleagues (2021) identified two commitment preservation mechanisms that could explain why prospective students stick to the program of their original choice. Especially the so-called self-fixation mechanism might be relevant in this context. This mechanism explains how students may deliberately downplay contradicting information. In other words, if the matching advise makes clear that a student's interests do not align with the program as much as they thought, they may ignore the advice (i.e., downplaying the mismatch) and start searching for components of the program that do match with their interests.

It would be useful to investigate whether the way of advising students as a result of matching can be improved. The system of providing concrete advice (i.e., traffic light analogies or positive-negative systems) may be abandoned in favor of a system of providing feedback. For example, Utrecht University does not provide advice, but generic feedback through which they hope to spark reflection regarding program choice. Since the matching procedures at Utrecht University are considered effective with respect to the three criteria for evaluating effectiveness posed in this dissertation, advising students might not be necessary for a successful matching procedure. However, it may be worthwhile to do further research into ways to send across the message of a perceived mismatch in such a way that prospective students accept the message and reflect on it (i.e., without directly telling a student that they should opt for another program). Neimeijer (2020) concludes that matching feedback should meet certain requirements on both content and form to be considered by prospective students. She poses that the content of the feedback should address current performance in relation to desired performance and how potential deficits can be addressed. The form of the feedback should be a concise text in an informal but serious tone, that is preferably supported by infographics. Such a feedback system can be especially attractive in the case of online matching procedures. However, feedback can also be incorporated in the emails most programs currently use for communicating their advice. In conclusion, in the way of advising students there is potential for further professionalizing the matching procedures and incorporating them in tutoring systems that are already in place in most universities. If a prospective student can discuss the received feedback with their future tutor, the tutor could help students reflecting on their choice and tutors will be instantly aware which students may need some extra guidance in the transitioning phase.

6.6 Limitations and Future Research

There are several limitations to the research presented in this dissertation. First, a problem with academic outcome data lies in the fact that these measures are generally skewed (i.e., number of earned credits) or imbalanced (i.e., percentage of dropouts). In chapter 4 we dealt with this issue by incorporating first-year GPA as a second outcome measure on top of earned credits, since GPA is regarded as a highly reliable measure (Bacon & Bean, 2006; Beatty et al., 2015). In Chapter 5, we applied undersampling and oversampling techniques, and reported weighted outcomes to reduce the negative impact of our imbalanced data on the findings (Dalipi et al., 2018). However, both in the case of credits and dropout, results should be interpreted with some caution.

The studies conducted in Chapter 2 and Chapter 3 provided indications that matching procedures, and especially sense of belonging, might play a different role for STEM students than for non-STEM students. Based on these results, we hypothesized that online matching procedures might serve the needs of STEM students better than matching days, while the opposite would apply to non-STEM students. We were about to study this, when the global Covid-19 pandemic emerged. We had to drop the research for a lack of on campus matching procedures to serve as reference groups to the online procedures. We incorporated the differential prediction in our study in Chapter 4 and indeed found differences in significance and strength of predictors of earned credits and first-year GPA between STEM and non-STEM students. Therefore, we conclude that it is worth further exploring the effects of different types of matching procedures in general and on the role of sense of belonging, in relation to student experiences, enrolment behavior, and academic outcomes across different fields of study.

Differential prediction by field of study is not the only aspect that should be further explored. It is advisable that the Netherlands Initiative for Education Research (NRO) incorporates detailed data on matching procedures in the to-be-established National Cohort Research Higher Education (NCO HO). If such a dataset can be established, the relation between types of matching procedures and dropout can be investigated more concisely. On top of that, not only research on the types of matching procedures, but also on the way these types of procedures have taken shape can provide more insight into what works and what does not. Procedures that are similar on paper, may be very different in practice. For example, regarding the matching days, we have not been able to differentiate between programs that required their students to prepare homework or not, and to make a test or not. Many more of such differentiations in practice could be studied to further our understanding of the way in which matching procedures impact students' program choice and dropout decisions. Lastly, as already touched upon in the practical implications, it would be worthwhile to research how the advice procedure could be optimized and potentially be integrated in existing support mechanisms.

A final limitation of the research in this study is the generalizability to the wider higher education context. The Netherlands has a stratified educational system, resulting in a relatively homogenous group of students regarding general cognitive ability (e.g., Resing & Drenth, 2007). Moreover, there is a wide array of open-admissions programs in Dutch higher education. That means that standardized tests used in other countries, such as the Scholastic Aptitude Test (SAT) and American College Test (ACT) in the USA, are generally not useful in Dutch admission procedures be it in the context of matching or selection. Likewise, matching procedures may not work in countries with less stratified educational systems. Institutions or policy makers in other countries wishing to implement similar matching procedures should critically consider the educational context in which they wish to apply the findings in this dissertation.

If other countries wish to implement similar procedures into their higher education admission procedures, it is important to thoroughly think through the infrastructure that should accompany it, prior to the implementation of the procedures. Especially if it should be possible to evaluate the effectiveness of these procedures several years after their implementation. The most adequate method for such an evaluation is a quasi-experiment, in which a representative group of students is subjected to matching procedures while other students follow the standard admission procedure. Potential problems that may be anticipated in the implementation of matching procedures concern differences in university data registration systems and the extent to which historical data are saved. If all these aspects are considered, very rich data can be acquired for comparison of student perceptions of the admission procedures, enrolment behavior, student wellbeing during the first year, and academic success. Based on the results of these comparisons, matching procedures could then be adjusted and implemented for the whole student population.

6.7 Conclusion

One of the major challenges in higher education research is reducing student dropout. We aimed to explain how matching procedures, which were implemented for improving program choice and student retention, can contribute to that goal. An important result is that prospective students find these matching procedures more useful if more elements of person-environment fit (i.e., ability beliefs, interests, and sense of belonging) can be tested. Moreover, it is important that these procedures provide students with a realistic view of the program of their choice. Matching procedures that are experienced as allowing for testing of more elements of fit, are also the ones that are associated with lower enrolment rates. This may entail that the more a matching procedure allows for testing fit, the more students may choose not to enrol when they experience a lack of fit. Furthermore, we have shown that pre-university indicators of fit are associated with first-year academic success and that certain indicators (e.g., high-school GPA) might be stronger for STEM students, while other indicators (e.g., non-cognitive traits like conscientiousness) may be stronger for non-STEM students. Lastly, it seems promising to use text mining techniques for predicting dropout from motivation texts.

References

Abraham, C., Richardson, M., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta analysis. *Psychological Bulletin, 138*(2), 353-387. https://doi.org/ 10.1037/a0026838

Acevedo, J. M., & Yancey, G. B. (2011). Assessing new employee orientation programs. *Journal of Workplace Learning, 23*(5), 349–354. https://doi.org/10.1108/13665621111141939

Aggarwal, C. C., & Zhai, C. (2012). A survey of text classification algorithms. In C. C., Aggarwal & C. Zhai (Eds.), *Mining text data* (pp. 163-222). Springer. https://doi.org/ 10.1007/978-1-4614

Alarcon, G. M., & Edwards, J. M. (2013). Ability and motivation: Assessing individual factors that contribute to university retention. *Journal of Educational Psychology 105*(1), 129-137. https://doi.org/10.1037/a0028496

Astin, A. W. (1993). *What Matters in College: Four Critical Years Revisited*. Jossey-Bass.

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*(6), 359-372. https://doi.org/10.1037/h0043445

Bacon, D. R., & Bean, B. (2006). GPA in Research Studies: An Invaluable but Neglected Opportunity. *Journal of Marketing Education, 28*(1), 35-42. https://doi.org/10.1177/0273475305284638

Bandura, A. (1977). Self-efficacy: towards a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Barrow, L., Sartain, L., & de la Torre, M. (2018). Selective enrolment high schools in Chicago: Admission and impacts. *University of Chicago Consortium on School Research*.

Bean, J. P. (2005). Nine themes of college student retention. In A. Seidman (Ed.), *College student retention: Formula for student success* (pp. 215-244). Praeger Publishers.

Beatty, A. S., Walmsley, P. T., Sackett, P. R., Kuncel, N. R., & Koch, A. J. (2015). The reliability of college grades. *Educational Measurement: Issues and Practice, 34*(4), 31-40. https://doi.org/10.1111/emip.12096

Beaulac, C., & Rosenthal, J. S. (2019). Predicting university students' academic success and major using random forests. *Research in Higher Education, 60*(7), 1048-1064. https://doi.org/10.1007/s11162-019-09546-y

Betts, J.R., & Morell, D. (1999). The determinants of undergraduate grade point average: the relative importance of family background, high school resources, and peer group effects. *The Journal of Human Resources, 34.* https://doi.org/10.2307/146346

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993-1022.

Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociological Methods & Research, 17*(3), 303-316. https://doi.org/10.1177/0049124189017003004

Bong, M. (2001). Role of self-efficacy and task-value in predicting college students' course performance and future enrolment intentions. *Contemporary Educational Psychology, 26*(4), 553-570. https://doi.org/10.1006/ceps.2000.1048

R

Boot, P., Zijlstra, H., & Geenen, R. (2017). The Dutch translation of the Linguistic Inquiry and Word Count (LIWC) 2007 dictionary. *Dutch Journal of Applied Linguistics, 6*(1), 65-76.

Borghans, L., Golsteyn, B. H., Heckman, J., & Humphries, J. E. (2011). Identification problems in personality psychology. *Personality and Individual Differences, 51*(3), 315-320. https://doi.org/10.1016/j.paid.2011.03.029

Bridgeman, B. (2013). Human ratings and automated essay evaluation. In M. D. Shermis and J. C. Burstein (Eds.), *Handbook of Automated Essay Evaluation: Current Applications and New Directions*, (pp. 221-232). Routledge.

Brinkworth, R., McCann, B., Matthews, C., & Nordström, K. (2009). First year expectations and experiences: Student and teacher perspectives. *Higher Education, 58*(2), 157-173. https://doi.org/10.1007/s/10734-008-9188-3

Brown, S. D., Lent, R. W., & Larkin, K. C. (1989). Self-efficacy as a moderator of scholastic aptitude: Academic performance relationships. *Journal of Vocational Behavior, 35*(1), 64-75. https://doi.org/10.1016/0001-8791(89)90048-1

Brown, S. D., Tramayne, S., Hoxha, D., Telander, K., Fan, X., & Lent, R. W. (2008). Social cognitive predictors of college students' academic performance and persistence: A meta-analytic path analysis. *Journal of Vocational Behavior, 72*(3), 298-308. https://doi.org/10.1016/j.jvb.2007.09.003

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual differences, 29*(6), 1057-1068. https://doi.org/10.1016/S0191-8869(99)00253-6

Camara, W. J., & Echternacht, G. (2000). The SAT [R] I and high school grades: Utility in predicting success in college. Research Notes.

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. 2, pp. 687–732). Consulting Psychologist Press.

Carvalho, R. G. G. (2016). Gender differences in academic achievement: The mediating role of personality. *Personality and Individual Differences, 94*, 54-58. https://doi.org/10.1016/j.paid.2016.01.011

Clifton, R. A., Perry, R. P., Roberts, L. W., & Peter, T. (2008). Gender, psychosocial dispositions, and the academic achievement of college students. *Research in Higher Education, 49*(8), 684-703. https://doi.org/10.1007/s11162-008-9104-9

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Lawrence Erlbaum Associates.

Conard, M. A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality, 40*(3), 339-346. https://doi.org/10.1016/j.jrp.2004.10.003

Conger, D., & Long, M. C. (2010). Why are men falling behind? Gender gaps in college performance and persistence. *The ANNALS of the American Academy of Political and Social Science, 627,* 184-214. https://doi.org/10.1177/0002716209348751

Crisp, G., Palmer, E., Turnbull, D., Nettelbeck, T., Ward, L., LeCouteur, A., Sarria, A., Strelan, P., & Schneider, L. (2009). First year student expectations: Results from a university wide survey. *Journal of University Teaching and Learning Practice, 6*(1).

Crossley, S., Paquette, L., Dascalu, M., McNamara, D. S., & Baker, R. S. (2016). Combining click-stream data with NLP tools to better understand MOOC completion. *Proceedings of the sixth international conference on learning analytics & knowledge*, 6-14.

Crosta, P. M. (2013). Characteristics of Early Community College Dropouts. CCRC Analytics. *Community College Research Center, Columbia University*.

Dalipi, F., Imran, A. S., & Kastrati, Z. (2018). MOOC dropout prediction using machine learning techniques: Review and research challenges. *2018 IEEE Global Engineering Education Conference (EDUCON),* 1007-1014.

Dana, J., Dawes, R., & Peterson, N. (2013). Belief in the unstructured interview: The persistence of an illusion. *Judgement and Decision Making, 8*(5), 512-520.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* Plenum.

Deci, E.L., & Ryan, R.M. (2008). Self-Determination Theory: A macrotheory of human motivation, development, and health. *Canadian Psychology, 49*(3), 182-185. https://doi.org/10.1037/a0012801

Delen, D. (2010). A comparative analysis of machine learning techniques for student retention management. *Decision Support Systems, 49*(4), 498-506. https://doi.org/10.1016/j.dss.2010.06.003

Denissen, J. J., Geenen, R., Van Aken, M. A., Gosling, S. D., & Potter, J. (2008). Development and validation of a Dutch translation of the Big Five Inventory (BFI). *Journal of personality assessment, 90*(2), 152-157. https://doi.org/10.1080/00223890701845229

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35*(1), 13-21. https://doi.org/10.1016/j.intell.2006.02.001

De Fruyt, F., & Mervielde, I. (1996). Personality and interests as predictors of educational streaming and achievement. *European Journal of Personality, 10*(5), 405-425. https://doi.org/10.1002/(SICI)1099-0984(199612)10:5<405::AID-PER255>3.0.CO;2-M

De Rome, E., & Lewin, T. (1984). Predicting persistence at university from information obtained at intake. *Higher Education, 13*(1), 49-66. https://doi.org/10.1007/BF00136530.

Dumfart, B., & Neubauer, A. C. (2016). Conscientiousness is the most powerful noncognitive predictor of school achievement in adolescents. *Journal of individual Differences, 37*(1), 8-15. https://doi.org/10.1027/1614-0001/A000182

Eccles, J., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J., and Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence, (Ed.), *Achievement and Achievement Motives* (pp. 75-146). Freeman.

Eccles, J.S., (2005). Subjective task value and the Eccles et al. model of achievement-related choices. In A.J. Elliot, C.S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 105-121). The Guilford Press.

Eccles, J.S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology 53*(1), 109-132. https:doi.org/10.1146/annurev.psych.53.100901.135153

Eisenberg, N., Duckworth, A. L., Spinrad, T. L., & Valiente, C. (2014). Conscientiousness: Origins in childhood? *Developmental psychology, 50*(5), 1331-1349. https://doi.org/10.1037/a0030977

EMBO. (2006). *From school to university – a report on the transition from secondary school biology education to university in Europe*. http://www.anisn.it/matita_allegati/pdf/report%20EMBO%202006.pdf.

Etcheverry, E., Clifton, R. A., & Roberts, L. W. (2001). Social capital and educational attainment: A study of undergraduates in a faculty of education. *Alberta Journal of Educational Research, 47*(1), 24-39. https://doi.org/10.11575/ajer.v47i1.54841

European Commission/EACEA/Eurydice (2015). *The European Higher Education Area in 2015: Bologna Process Implementation Report.* Luxembourg: Publications Office of the European Union. http://www.eurostudent.eu/download_files/documents/2015_Implementation_report_20.05.2015.pdf

Farsides, T., & Woodfield, R. (2007). Individual and gender differences in good and first-class undergraduate degree performance. *British Journal of Psychology, 98*(3), 467-483. https://doi.org/10.1348/000712606X150246

Feldman, K. A. Smart, J. C., and Ethington, C. (1999). Major field and Person-Environment Fit. *The Journal of Higher Education, 70*(6), 642-669. https://doi.org/10.1080/00221546.1999.11780802

Fitz-Walter, Z., Wyeth, P., Tjondronegoro, D., & Johnson, D. (2014). Exploring the effect of achievements on students attending university orientation. *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play,* 87-96.

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. https://doi.org/10.1016/j.lindif.2017.05.003

Freeman, T. M., Anderman, L. H., and Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education, 75*(3), 203-220. https://doi.org/10.3200/JEXE.75.3.203-220

Gigliotti, R. J., & Huff, H. K. (1995). Role-related conflicts, strains and stresses of older-adult college students. *Sociological Focus, 28*(3), 329-342. https://doi.org/10.1080/00380237.1995.10571057

Goldberg, L. R. (1990). An alternative "description of personality:" The Big-Five factor structure. *Journal of Personality and Social Psychology, 59*(6), 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216

Goovaerts, M. (2012). Wie overleeft het eerste bachelor-jaar niet? Een onderzoek naar drop-out in het hoger onderwijs. [Master Thesis, Antwerp University].

Green, T. (2016). A methodological review of structural equation modelling in higher education research. *Studies in Higher Education, 41*(12), 2125-2155. https://doi.org/10.1080/03075079.2015.1021670

Guiffrida, D. A., Lynch, M. F., Wall, A. F., & Abel, D. S. (2013). Do reasons for attending college affect academic outcomes?: A test of a motivational model from a self-determination theory perspective. *Journal of College Student Development, 54*(2), 121-139. https://doi.org/10.1353/csd.2013.0019

Hällsten, M. (2017). Is education a risky investment? The scarring effect of university dropout in Sweden. *European Sociological Review, 33*(2), 169-181. https://doi.org/10.1093/esr/jcw053

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. https://doi.org/10.1037/0022-0663.94.3.562

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology, 100*(1), 105-122. https://doi.org/10.1037/0022-0663.100.1.105

Heublein, U., Hutzsch, C., Schreiber, J., Sommer, D., & Besuch, G. (2010). Ursachen des Studienabbruchs in Bachelor- und in herkömmlichen Studiengängen: *Ergebnisse einer bundesweiten Befragung von Exmatrikulierten des Studienjahres, 8*(2).

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127. https://doi.org/10.1207/s15326985ep4102_4

Hofer, M. (2010). Adolescents' development of individual interests: A product of multiple goal regulation? *Educational Psychologist, 45*(3), 149–166. https://doi.org/10.1080/00461520.2010.493469.

Hoffman, M., Richmond, J., Morrow, J., & Salomone, K. (2002). Investigating "sense of belonging" in first-year college students. *Journal of College Student Retention: Research, Theory & Practice, 4*(3), 227-256. https://doi.org/10.2190/DRYC-CXQ9-JQ8V-HT4V

Hofman, A., & Van Den Berg, M. (2000). Determinants of study progress: The impact of student, curricular, and contextual factors on study progress in university education. *Higher Education in Europe, 25*(1), 93-110. https://doi.org/10.1080/03797720050002242

Holmegaard, H. T., Madsen, L. M., & Ulriksen, L. (2016). Where is the engineering I applied for? A longitudinal study of students' transition into higher education engineering, and their considerations of staying or leaving. *European Journal of Engineering Education, 41*(2), 154-171. https://doi.org/10.1080/03043797.2015.1056094

Holmegaard, H. T., Ulriksen, L. M., & Madsen, L. M. (2015). A narrative approach to understand students' identities and choices. In E. K. Henriksen, J. Dillon, & J. Ryder (Eds.), *Understanding student participation and choice in science and technology education* (pp. 31–42). Springer Science+Business Media Education.

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*, 63-84. https://doi.org/10.1016/j.edurev.2015.11.002

Hough, L. M., & Schneider, R. J. (1996). Personality traits, taxonomies, and applications in organizations. *Individual differences and Behavior in Organizations,* 31-88. https://doi.org/10.1111/j.1754-9434.2008.00048.x

Hutt, S., Gardener, M., Kamentz, D., Duckworth, A. L., & D'Mello, S. K. (2018). Prospectively predicting 4-year college graduation from student applications. *Proceedings of the eighth international conference on learning analytics and knowledge,* 280-289.

Inspectie van het Onderwijs. (2019). Internationalisering en de toegankelijkheid van het Nederlands hoger onderwijs.

International Educational Data Mining Society. (n.d.). https://www.educationaldatamining.org

Jaeger, D. A., & Page, M. E. (1996). Degrees matter: New evidence on sheepskin effects in the returns to education. *The Review of Economics and Statistics, 78*(4), 733-740. https://doi.org/10.2307/2109960

Jansen, E. P. (2004). The influence of the curriculum organization on study progress in higher education. *Higher education, 47*(4), 411-435. https://doi.org/10.1023/B:HIGH.0000020868.39084.21

Jayaraman, J. (2020). Predicting Student Dropout by Mining Advisor Notes. *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, 629-632.

R

John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138). Guilford Press.

Johnes, G., & McNabb, R. (2004). Never give up on the good times: Student attrition in the UK. *Oxford Bulletin of Economics and Statistics, 66*(1), 23-47. https://doi.org/10.1111/j.1468-0084.2004.00068.x

Jongbloed, B. W. A., de Boer, H. F., Kaiser, F., & Vossensteyn, H. J. J. (2018). *Bekostiging van het Nederlandse hoger onderwijs: kostendeterminanten en varianten*. Center for Higher Education Policy Studies (CHEPS).

Kappe, R., & van der Flier, H. (2010). Using multiple and specific criteria to assess the predictive validity of the Big Five personality factors on academic performance. *Journal of Research in Personality, 44*(1), 142-145. https://doi.org/10.1016/j.jrp.2009.11.002

Kappe, R., & van der Flier, H. (2012). Predicting academic success in higher education: what's more important than being smart? *European Journal of Psychology of Education, 27*(4), 605-619. https://doi.org/10.1007/s10212-011-0099-9

Keith, T. Z. 2006. *Multiple Regression and Beyond*. Pearson.

Keup, J. R., & Stolzenberg, E. B. (2004). *The 2003 your first college year (YFCY) survey: Exploring the academic and personal experiences of first-year students* (No. 40). National Resource Center for the First-Year Experience & Students in Transition.

Kira Talent (2018). *Breaking down bias in admissions: The how-to guide to preventing admissions bias at your school.* http://start.kiratalent.com/breaking-down-admissions-bias/.

Kirk, G. (2018). Retention in a Bachelor of Education (Early childhood studies) course: Students say why they stay and others leave. *Higher Education Research & Development, 37*(4), 773-787. https://doi.org/10.1080/07294360.2018.1455645

Kline, R.B. (2011). *Principles and Practice of Structural Equation Modelling.* Guilford Publications.

Kling, K., Noftle, E. E., & Robins, R. W. (2012). Why do standardized tests underpredict women's academic performance? The role of conscientiousness. *Social Psychological and Personality Science, 4*(5), 600-606. https://doi.org/10.1177/1948550612469038

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences, 51*(4), 472-477. https://doi.org/10.1016/j.paid.2011.04.019

Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text classification algorithms: A survey. *Information, 10*(4), 150-218. https://doi.org/10.3390/info10040150

Krause, K. L., & Coates, H. (2008). Students' engagement in first-year university. *Assessment & Evaluation in Higher Education, 33*(5), 493-505. https://doi.org/10.1080/02602930701698892

Kristof-Brown, A., & Guay, R. P. in S. Zedeck (Ed). (2011). *APA handbook of industrial and organizational psychology,* Vol 3: Maintaining, expanding, and contracting the organization., (pp. 3-50). American Psychological Association, viii, 960 pp. https://doi.org/10.1037/12171-001

Kucel, A., & Vilalta-Bufí, M. (2013). Job satisfaction of university graduates. *Revista de Economía Aplicada, 21*(61), 29-55.

Kurysheva, A., van Rijen, H. V., & Dilaver, G. (2019). How do admission committees select? Do applicants know how they select? Selection criteria and transparency at a Dutch University. *Tertiary Education and Management, 25*(4), 367-388. https://doi.org/10.1007/s11233-019-09050-z

Lassibille, G., & Navarro Gómez, L. (2008). Why do higher education students drop out? Evidence from Spain. *Education Economics, 16*(1), 89-105. https://doi.org/10.1080/09645290701523267

Lauría, E. J., Baron, J. D., Devireddy, M., Sundararaju, V., & Jayaprakash, S. M. (2012). Mining academic data to improve college student retention: An open source perspective. *Proceedings of the Second International Conference on Learning Analytics and Knowledge*, 139-142.

Leenheer, J. (2022). Do you think we are a match? The predictive power of matching activities for prospective students of an international business program. *The International Journal of Management Education, 20*(2), 100637. https://doi.org/10.1016/j.ijme.2022.100637

Lent, R. W., Brown, S. D., & Larkin, K. C. (1984). Relation of self-efficacy expectations to academic achievement and persistence. *Journal of Counseling Psychology, 31*(3), 356-362. https://doi.org/10.1037/0022-0167.31.3.356

Lent, R. W., Brown, S. D., & Larkin, K. C. (1987). Comparison of three theoretically derived variables in predicting career and academic behavior: Self-efficacy, interest congruence, and consequence thinking. *Journal of Counseling Psychology, 34*(3), 293-298. https://doi.org/10.1037/0022-0167.34.3.293

Lewin, K. (1935). A Dynamic Theory of Personality. *Selected Papers.* Graw-Hill Book Company.

Mattanah, J. F., Ayers, J. F., Brand, B. L., Brooks, L. J., Quimby, J. L., & McNary, S. W. (2010). A social support intervention to ease the college transition: Exploring main effects and moderators. *Journal of College Student Development, 51*(1), 93-108. https://doi.org/10.1353/csd.0.0116

McFarlane, K. (2018). Higher education learner identity for successful student transitions, *Higher Education Research & Development, 37*(6), 1201-1215. https://doi.org/10.1080/07294360.2018.1477742

Meens, E.E.M. (2018). Motivation: Individual differences in students' educational choices and study success. [Doctoral dissertation, Tilburg University].

Meeuwisse, M., Severiens, S. E., & Born, M. Ph. (2010). Reasons for withdrawal from higher vocational education. A comparison of ethnic minority and majority non-completers, *Studies in Higher Education, 35*(1), 93-111. https://doi.org/10.1080/03075070902906780

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation.* Jossey-Bass.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30-38. https://doi.org/10.1037/0022-0167.38.1.30

Munro, B. H. (1981). Dropouts from higher education: Path analysis of a national sample. *American Educational Research Journal, 18*(2), 133-141. https://doi.org/10.2307/1162377

Murphy, C., Hawkes, L., & Law, J. (2002). How international students can benefit from a web-based college orientation. *New Directions for Higher Education, 117*, 37–44. https://doi.org/10.1002/he.45.

Muthén, B.O., & Muthén, L.K. (2019). MPlus (8.3).

Naylor, R., Baik, C., & Arkoudis, S. (2018). Identifying attrition risk based on the first year experience. *Higher Education Research & Development, 37*(2), 328-342. https://doi.org/10.1080/07294360.2017.1370438

Nederlands Jeugdinstituut. (2021). *Cijfers over Onderwijsprestaties.* https://www.nji.nl/cijfers/onderwijsprestaties.

Neimeijer, I. (2020). Bruikbare Feedback op een Studiekeuzetest: Ontwerpgericht Onderzoek in het Hoger Onderwijs. [Master Thesis, Open University].

R

Nguyen, N., Muilu, T., Dirin, A., & Alamäki, A. (2018). An interactive and augmented learning concept for orientation week in higher education. *International Journal of Educational Technology in Higher Education, 15*(1), 1-15. https://doi.org/10.1186/s41239-018-0118-x

Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2016). Predicting performance in higher education using proximal predictors. *PloS ONE, 11*(4). https://doi.org/10.1371/journal.pone.0153663

Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2017). Measuring non-cognitive predictors in high-stakes contexts: The effect of self-presentation on self-report instruments used in admission to higher education. *Personality and Individual Differences, 106,* 183–189. https://doi.org/10.1016/j.paid.2016.11.014.

Niessen, A. S. M. (2017). New Rules, New Tools: Predicting Academic Achievement in College Admissions [Doctoral Dissertation, University of Groningen].

Niessen, A. S. M., & Meijer, R. R. (2017). On the use of broadened admission criteria in higher education. *Perspectives on Psychological Science, 12*(3), 436-448. https://doi.org/10.1177/1745691616683050

Niessen, A. S. M. & Meijer, R. R. (2020). Character-based admissions criteria in the United States and in Europe: Rationale, evidence, and some critical remarks. In M.E. Oliveri & C. Wendler (Eds.) *Higher Education Admission Practices: An International Perspective* (pp. 76-95). Cambridge University Press. https://doi.org/10.1017/9781108559607.005

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*(5), 971-990. https://doi.org/10.1016/j.paid.2007.03.017

O'Keefe, M., Laven, G., & Burgess, T. (2011). Student non-completion of an undergraduate degree: Wrong program selection or part of a career plan? *Higher Education Research & Development, 30*(2), 165-177. https://doi.org/10.1080/07294360.2010.512630

Oreopoulos, P., & Petronijevic, U. (2013) *Making college worth it: A review of research on the returns to higher education.* NBER Working Paper Series.

Pajares, F. (1996). Self-efficacy beliefs and mathematical problem-solving of gifted students. *Contemporary Educational Psychology, 21*(4), 325-344. https://doi.org/10.1006/ceps.1996.0025

Pajares, F. (1997). Current directions in self-efficacy research. *Advances in Motivation and Achievement, 10*(149), 1-49.

Pampaka, M., Williams, J., Hutcheson, G., Wake, G., Black, L., Davis, P., & Hernandez-Martinez, P. (2012). The association between mathematics pedagogy and learners' dispositions for university study. *British Educational Research Journal, 38*(3), 473-496. https://doi.org/10.1080/01411926.2011.555518

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. (2007). The development and psychometric properties of LIWC2007 (pp. 1–22). Pennebaker Conglomerates.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385-407.
http://dx.doi.org/10.1007/s10648-004-0006-x

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33. https://doi.org/10.1037/0022-0663.82.1.33

Pitman, T. (2016). Understanding 'fairness' in student selection: are there differences and does it make a difference anyway? *Studies in Higher Education, 41*(7), 1203- 1216.
https://doi.org/10.1080/03075079.2014.968545

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322-338. https://doi.org/10.1037/a0014996

Porter, S. R., & Umbach, P. D. (2006). College major choice: An analysis of person–environment fit. *Research in Higher Education, 47*(4), 429-449. https://doi.org/10.1007/s11162-005-9002-3

Posselt, J. R. (2016). *Inside graduate admissions: Merit, diversity, and faculty gatekeeping.* Harvard University Press.

Psacharopoulos, G. (1994). Returns to investment in education: A global update. *World Development, 22*(9), 1325-1343. https://doi.org/10.1016/0305-750X(94)90007-8

Rausch, J. L., & Hamilton, M. W. (2006). Goals and distractions: explanations of early attrition from traditional university freshmen. *The Qualitative Report, 11*(2), 317-334. https://doi.org/10.46743/2160-3715/2006.1676

Renninger, K. A., & Hidi, S. (2017). *The power of interest for motivation and engagement*. Routledge. https://doi.org/10.4324/9781315771045

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. *Encyclopedia of database systems*, *5*, 532-538. https://doi.org/10.1007/978-1-4899-7993-3_565-2

Reumer, C. & Van der Wende, M. (2010). *Excellence and Diversity : The emergence of selective admission policies in Dutch higher education – a case study on Amsterdam University College.* Research & Occasional Paper Series: CSHE.15.10

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

Sahlgren, M. (2008). The distributional hypothesis. *Italian Journal of Disability Studies*, *20*, 33-53.

Samuel, R., & Burger, K. (2020). Negative life events, self-efficacy, and social support: Risk and protective factors for school dropout intentions and dropout. *Journal of Educational Psychology, 112*(5), 973-987. https://doi.org/10.1037/edu0000406

Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., De Fruyt, F., & Duyck, W. (2019). The effects of vocational interest on study results: Student person-environment fit and program interest diversity. *PLoS ONE, 14*(4). https://doi.org/10.1371/journal.pone.0214618

Schunk, D. H. (1995). Self-efficacy, motivation, and performance. *Journal of Applied Sport Psychology, 7*(2), 112-137. https://doi.org/10.1080/10413209508406961

Schunk, D. H., & Pajares, F. (2009). Self-Efficacy Theory. In: *Handbook of motivation at school* (pp. 49-68). Routledge.

Sheard, M. (2009). Hardiness commitment, gender, and age differentiate university academic performance. *British Journal of Educational Psychology, 79*(1), 189-204. https://doi.org/10.1348/000709908X304406

Sneyers, E., & De Witte, K. (2018). Interventions in higher education and their effect on student success: a meta-analysis. *Educational Review*, *70*(2), 208-228. https://doi.org/10.1080/00131911.2017.1300874

Soppe, K.F.B., Wubbels, T., Leplaa, H.J., Klugkist, I.G., & Wijngaards-de Meij, L.D.N.V. Do they match? Prospective students' experiences with choosing university programmes, *European Journal of Higher Education,* 9(4), 359-376. https://doi.org/10.1080/21568235.2019.1650088

Statistics Netherlands. (2017a). *Wo-cohorten; eerste wo-diploma, studierichting*. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83285NED/table

R

Statistics Netherlands. (2017b). *Wo-cohorten; hoogst behaalde wo-diploma, vooropleiding*. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83305NED/table?ts=1517826792661

Steenman, S. C., Bakker, W. E., & van Tartwijk, J. W. (2016). Predicting different grades in different ways for selective admission: disentangling the first-year grade point average. *Studies in Higher Education, 41*(8), 1408-1423. https://doi.org/10.1080/03075079.2014.970631.

Stone, C., Quirk, A., Gardener, M., Hutt, S., Duckworth, A. L., & D'Mello, S. K. (2019). Language as thought: Using natural language processing to model noncognitive traits that predict college success. *Proceedings of the Ninth International Conference on Learning Analytics & Knowledge,* 320-329.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273–285). Sage Publications Ltd.

Strayhorn, T.L., (2012). *College students' sense of belonging, a key to educational success for all students.* Routledge.

Swenson, L. M., Nordstrom, A., & Hiester, M. (2008). The role of peer relationships in adjustment to college. *Journal of College Student Development, 49*(6), 551-567. https://doi.org/10.1353/csd.0.0038

Thammasiri, D., Delen, D., Meesad, P., & Kasap, N. (2014). A critical assessment of imbalanced class distribution problem: The case of predicting freshmen student attrition. *Expert Systems with Applications, 41*(2), 321-330. https://doi.org/10.1016/j.eswa.2013.07.046

Ting, S. M. R., & Robinson, T. L. (1998). First-year academic success: A prediction combining cognitive and psychosocial variables for Caucasian and African American students. *Journal of College Student Development, 39*(6), 599-610.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45*(1), 89-125. https://doi.org/10.3102/00346543045001089

Tinto, V. (1987). *Leaving College: Rethinking the Causes and Cures of Student Attrition.* University of Chicago Press. https://doi.org/10.1080/00221546.1988.11780239

Tinto, V. (1993). *Leaving College: Rethinking the Causes and Cures of Student Attrition. Second Edition.* University of Chicago Press. https://doi.org/10.7208/chicago/9780226922461.001.0001

Tolstova, I.N. (2006). Between school and college: is the gap getting wider? Thoughts of a sociology instructor. *Russian Education & Society, 48*(6), 7-26. https://doi.org/10.2753/RES1060-9393480601

Trapmann, S., Hell, B., Hirn, J.-O. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Journal of Psychology, 215*(2), 132-151. https://doi.org/10.1027/0044-3409.215.2.132

Tross, S. A., Harper, J. P., Osher, L. W., & Kneidinger, L. M. (2000). Not just the usual cast of characteristics: Using personality to predict college performance and retention. *Journal of College Student Development, 41*(3), 323–334.

Trowler, V. (2010). Student engagement literature review. *The Higher Education Academy*, *11*(1), 1-15.

Ulriksen, L., Madsen, L. M. & Holmegaard, H. T. (2010) What do we know about explanations for drop out/opt out among young people from STM higher education programmes? *Studies in Science Education, 46*(2), 209-244. https://doi.org/10.1080/03057267.2010.504549

Umarji, O., McPartlan, P. & Eccles, J. (2018). Patterns of math and English self-concepts as motivation for college major selection. *Contemporary Educational Psychology, 53,* 146-158. https://doi.org/10.1016/j.cedpsych.2018.03.004

Unger, M., Wroblewski, A., Latcheva, R., Zaussinger, S., Hofmann, J. & Musik, C. (2009). Frühe Studienabbrüche an Universitäten in Österreich, Institut für Höhere Studien (IHS), Wien.

Van Gurp, W., & Van den Hurk, M. (2014). De relatie tussen schriftelijke feedback, feedback-perceptie en teacher-efficacy bij leraren in opleiding. *Tijdschrift voor Lerarenopleiders, 35*(1), 59-69.

Van den Broek, A., Warps, J., Cuppen, J., Termorshuizen, T., & Lodewick, J. (2020). *Monitor Beleidsmaatregelen Hoger Onderwijs 2019-2020.* Onderzoek in opdracht van het ministerie van OCW. ResearchNed.

Van Herpen, S.G.A., Meeuwisse, M., Hofman, W.H.A., Severiens, S.E. & Arends, L. R. (2017). Early predictors of first-year academic success at university: Pre-university effort, pre-university self-efficacy, and pre-university reasons for attending university. *Educational Research and Evaluation, 23*(1-2), 52-72. https://doi.org/10.1080/13803611.2017.1301261

Van der Veen, I., De Jong, U., Van Leeuwen, M., & Korteweg, J. A. (2005). The development of higher education students' interest in their subject: the case of higher professional education in the Netherlands. *Studies in Higher Education, 30*(3), 275-289. https://doi.org/10.1080/03075070500095705

Vansteenkiste, M., Ryan, R. M., & Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and emotion*, *44*(1), 1-31. https://doi.org/10.1007/s11031-019-09818-1

Veerman, C. P., Berdahl, R. M., Bormans, M. J. G., Geven, K. M., Hazelkorn, E., Rinnooy Kan, A. H. G., ... & Vossensteyn, J. J. (2010). Differentiëren in drievoud: omwille van kwaliteit en verscheidenheid van het hoger onderwijs: Advies van de Commissie Toekomstbestendig Hoger Onderwijs Stelsel.

Vereniging van Nederlandse Universiteiten [VSNU]. (n.d.(a)). *Bachelorfase status na 1 jaar.* https://www.universiteitenvannederland.nl/bachelorfase-status-na-1-jaar.html

Vereniging van Nederlandse Universiteiten [VSNU]. (n.d.(b)). *De studiekeuzecheck.* https://www.vsnu.nl/2014/studiekeuzecheck.html

Vereniging van Nederlandse Universiteiten [VSNU]. (2017). *De Studiekeuzecheck op 13 universiteiten.* https://www.universiteitenvannederland.nl/nl_NL/vsnu-publicaties.html

Vereniging van Nederlandse Universiteiten [VSNU]. (2016). *Status na 1 jaar: Doorstuderen en uitval studenten.* http://www.vsnu.nl/nl_NL/doorstuderen-en-uitval.html

Vossensteyn, J. J., Stensaker, B., Kottman, A., Hovdhaugen, E., Jongbloed, B., Wollscheid, S., Kaiser, F., & Cremonini, L. (2015). *Dropout and completion in Higher Education in Europe.* CHEPS.

Vulperhorst, J. P., Lutz, C., de Kleijn, R., & van Tartwijk, J. (2018). Disentangling the predictive validity of high school grades for academic success in university. *Assessment & Evaluation in Higher Education, 43*(3), 399-414. https://doi.org/10.1080/02602938.2017.1353586

Vulperhorst, J. P., Van der Rijst, M. R., & Akkerman, S. F. (2020). Dynamics in higher education choice: weighing one's multiple interests in light of available programmes. *Higher Education, 79*(6), 1001-1021. https://doi.org/10.1007/s10734-019-00452-x

Vulperhorst, J. P., Van der Rijst, R. M., Holmegaard, H. T., & Akkerman, S. F. (2021). Unravelling why students do or do not stay committed to a programme when making a higher education choice. *Journal of Further and Higher Education*, 1-16. https://doi.org/10.1080/0309877X.2021.1986686

R

Vulperhorst, J. P., Wessels, K. R., Bakker, A., & Akkerman, S. F. (2018). How do STEM-interested students pursue multiple interests in their higher educational choice? *International Journal of Science Education, 40*(8), 828–846. https://doi.org/10.1080/09500693.2018.1452306

Wang, M-T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychological Review, 29*(1), 119–140. https://doi.org/10.1007/s10648-015-9355-x

Ward, A. M., & Brennan, N. M. (2020) Developing a student doctoral education fit analytical model to assess performance. *Studies in Higher Education, 45*(7), 1448-1460. https://doi.org/10.1080/03075079.2018.1545758

Warps, J., Hogeling, L., Pass, J., & Brukx, D. (2009). *Studiekeuze en studiesucces: Een selectie van gegevens uit de Startmonitor over studiekeuze, studieuitval, en studiesucces in het hoger onderwijs*. Onderzoek in opdracht van SURF-Studiekeuze123. ResearchNed.

Warps, J., Nooij, J., Muskens, M., Kurver, B., & Broek, A. van den (2017). *De studiekeuzecheck. Landelijk onderzoek naar uitvoering en opbrengsten van de studiekeuzecheck in het hoger onderwijs*. Onderzoek in opdracht van het ministerie van OCW. ResearchNed.

Watson, G., Cavallaro Johnson, G., & Austin, H. (2004). Exploring relatedness to field of study as an indicator of student retention. *Higher Education Research & Development, 23*(1), 57-72. https://doi.org/10.1080/0729436032000168496

Wen, M., Yang, D., & Rose, C. (2014). Sentiment Analysis in MOOC Discussion Forums: What does it tell us? *Educational data mining 2014*.

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. R., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT scores, high school grades, and SES. *Educational Assessment, 20*(1), 23-45. https://doi.org/10.1080/10627197.2015.997614

Wet Kwaliteit in Verscheidenheid [Quality in Diversity Law] (2013). *Memorie van Toelichting [Note of Explanation]*. https://www.eerstekamer.nl/behandeling/20130118/memorie_van_toelichting_2/ document3/f=/vj6ifr0xnkyo.pdf

Wet op het Hoger onderwijs en Wetenschappelijk onderzoek (WHW) [Law on Higher Education and Scientific Research], article 7.8.

Whitney, D. R. (1969). Predicting from expressed vocational choice: A review. *The Personnel and Guidance Journal, 48*(4), 279–286. https://doi.org/10.1002/j.2164-4918.1969.tb03318.x

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*(1), 49-78. https://doi.org/10.1007/BF02209024

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*(3), 265-310. https://doi.org/10.1016/0273-2297(92)90011-P

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Willems, J., Coertjens, L., Tambuyzer, B., & Donche, V. (2019). Identifying science students at risk in the first year of higher education: The incremental value of non-cognitive variables in predicting early academic achievement. *European Journal of Psychology of Education*, 34(4), 847-872. https://doi.org/10.1007/s10212-018-0399-4

Willcoxson, L., & Wynder, M. (2010). The relationship between choice of major and career, experience of university and attrition. *Australian Journal of Education, 54*(2), 175-189. https://doi.org/10.1177/000494411005400205

Wlodkowski, R. J., Mauldin, J. E., & Gahn, S. W. (2001). Learning in the Fast Lane: Adult Learners' Persistence and Success in Accelerated College Programs. New Agenda Series [TM]. Volume 4, Number 1.

Wolter, S. C., Diem, A. & Messer, D. (2013). *Studienabbrüche an Schweizer Universitäten, Schweizerische Koordinationsstelle für Bildungsforschung*. SKBF.

Yorke, M. (2000). Smoothing the transition into higher education: What can be learned from student non-completion. *Journal of Institutional Research, 9*(1), 78–88.

Zajacova, A., Lynch, S. M. & Espenshade, T. J. (2005). Self-efficacy, stress and academic success in college. *Research in Higher Education, 46*(6), 677-706.
https://doi.org/10.1007/s11162-004-4139-z

Zhang, Y., Oussena, S., Clark, T., & Kim, H. (2010). Use Data Mining to Improve Student Retention in Higher Education-A Case Study. *ICEIS (1),* 190-197.

Zhang, Z. & RiCharde, S. (1998). Prediction and Analysis of Freshman Retention. 38th *Annual Conference of the Association for Institutional Research*.

Zheng, J. L., Saunders, K. P., Shelley II, M. C., & Whalen, D. F. (2002). Predictors of academic success for freshmen residence hall students. *Journal of College Student Development 43*(2), 267-283.

Zupanc, K. (2018). Semantics-based automated essay evaluation. [Doctoral dissertation, University of Ljubljana].

Zwick, R. (2017). Admissions tests and high school grades: what do they measure? In *Who Gets In?* (pp. 49-77). Harvard University Press. https://doi.org/10.4159/9780674977648-004

Zyphur, M. J., Bradley, J. C., Landis, R. S., & Thoresen, C. J. (2007). The effects of cognitive ability and conscientiousness on performance over time: A censored latent growth model. *Human Performance, 21*(1), 1-27. https://doi.org/10.1080/08959280701521967

Appendices

Appendix A. Factor analysis and measurement invariance for University 1

Exploratory factor analysis of Conscientiousness with factor loadings and reliability of the scale.

EFA 1-factor solution		
Conscientiousness	**Model 1**	**Model 2**
Does a thorough job	.549	.544
Perseveres until the task is finished	.549	.540
Tends to be disorganized[R]	.555	.563
Tends to be lazy[R]	.661	.676
Is a reliable worker	.487	.473
Does things efficiently	***.276***	
Makes plans and follows through with them	.569	.554
Is easily distracted[R]	.518	.514
Can be somewhat careless[R]	.577	.593
Reliability		Cronbach's α = .778

Note: Items with an R are reverse coded.

Establishing measurement invariance of the model by cohort

To determine whether data of the separate cohorts can be analyzed in one model, invariance across cohorts of the measurement model was established for a one-factor confirmatory factor analysis model with the latent variable conscientiousness. The initial configural model (without constraints) showed inadequate fit. Based on Modification Indices and theoretical reasons, three correlated residual terms were added (item 2 with 1 (perseveres x thorough job), item 5 with 2 (reliable worker x perseveres) and item 5 with 1 (reliable x thorough job)). Subsequently a model comparison test was specified to establish measurement invariance for the measurement model.

Traditionally, Chi-square tests are used for determining model fit. However, the Chi-Square statistic is highly sensitive to sample size, resulting in overly strict conclusions on model fit (Marsh et al., 1988). Hence, CFI was used as the main test statistic to determine model fit, as it is the most stable across different sample sizes (Cheung

& Rensvold, 2002). A cut-off point of ΔCFI ≤0.002 was used to determine the best fitting model (Meade et al., 2008). Decisions based on the CFI were supported by looking at the RMSEA, which is independent of sample size, but sensitive to model complexity (Van de Schoot et al., 2012), using a cut-off point of ΔRMSEA ≤0.007 (Meade et al., 2008) and AIC (Aikaike's Information Criterion), which makes a trade-off between fit and complexity of the model (Van de Schoot et al., 2012), with a lower value representing a better trade-off. For University 1 partial scalar invariance (equal factor loadings and partially equal intercepts) was established.

	University 1				
	CFI	ΔCFI	RMSEA	ΔRMSEA	AIC
Configural	0.945		0.065		182484.288
Metric	0.943	.002	0.058	.007	182479.824
Scalar	0.940	.003	0.055	.003	182494.709
Partial scalar[12]	**0.942**	**.001**[13]	**0.059**	**.001**	**182478.069**

Measurement invariance by discipline

To determine whether data of the separate disciplines can be analyzed in one model, invariance across disciplines of the measurement model was established for a one-factor confirmatory factor analysis model with the latent variable conscientiousness. The initial configural model (without constraints) showed inadequate fit. Based on Modification Indices and theoretical reasons, three correlated residual terms were added (item 2 with 1 (perseveres x thorough job), item 5 with 2 (reliable worker x perseveres) and item 5 with 1 (reliable x thorough job)). Partial scalar invariance was established.

	University 1				
	CFI	ΔCFI	RMSEA	ΔRMSEA	AIC
Configural	0.944		0.066		182406.544
Metric	0.943	.001	0.060	.006	182397.724
Scalar	0.937	.006	0.059	.001	182462.891
Partial scalar[14]	**0.941**	**.002**	**0.064**	**.004**	**182433.426**

[12]Items CS1 – CS5, CS7, and CS8 are constraint to be invariant, while error variance of CS9 is freed.

[13]To establish model fit of the partial scalar invariant model, it is compared to the metric invariant model, rather than the scalar invariant model.

[14]Items CS1 – CS5, CS7, and CS8 are constraint to be invariant, while error variance of CS9 is freed.

Appendix B. Factor analysis and measurement invariance for University 2

Exploratory factor analysis of Conscientiousness with factor loadings and reliability of the scale.

EFA 1-factor solution		
Conscientiousness	**Model 1**	**Model 2**
Does a thorough job	.550	.548
Perseveres until the task is finished	.532	.527
Tends to be disorganized[R]	.632	.638
Tends to be lazy[R]	.701	.712
Is a reliable worker	.517	.512
Does things efficiently	***.303***	
Makes plans and follows through with them	.601	.585
Is easily distracted[R]	.576	.573
Can be somewhat careless[R]	.569	.579
Reliability		Cronbach's α = .801

Note: Items with an R are reverse coded.

Establishing measurement invariance of the model by cohort.

To determine whether data of the separate cohorts can be analyzed in one model, invariance across cohorts of the measurement model was established for a one-factor confirmatory factor analysis model with the latent variable conscientiousness. The initial configural model (without constraints) showed inadequate fit. Based on Modification Indices and theoretical reasons, three correlated residual terms were added (item 5 with 2 (reliable worker x perseveres), item 9 with 3 (somewhat careless x disorganized) and item 4 with 1 (lazy x thorough job)). Subsequently a model comparison test was specified to establish measurement invariance for the measurement model. For University 2, full scalar invariance (equal factor loadings and equal intercepts) was established.

	University 2				
	CFI	ΔCFI	RMSEA	ΔRMSEA	AIC
Configural	0.943		0.072		273077.780
Metric	0.942	.001	0.066	.006	273080.564
Scalar	**0.941**	**.001**	**0.062**	**.004**	**273079.424**

Measurement invariance by discipline

To determine whether data of the separate disciplines can be analyzed in one model, invariance across disciplines of the measurement model was established for a one-factor confirmatory factor analysis model with the latent variable conscientiousness. The initial configural model (without constraints) showed inadequate fit. Based on Modification Indices and theoretical reasons, three correlated residual terms were added (item 5 with 2 (reliable worker x perseveres), item 9 with 3 (somewhat careless x disorganized) and item 4 with 1 (lazy x thorough job)). Partial scalar invariance was achieved.

	University 2				
	CFI	ΔCFI	RMSEA	ΔRMSEA	AIC
Configural	0.953		0.067		272628.121
Metric	0.951	.002	0.063	.004	272674.173
Scalar	0.946	.007	0.061	.002	272778.748
Partial scalar[15]	**0.950**	**.001**	**0.064**	**.003**	**272676.863**

[15]Items CS1 – CS5, CS7, and CS8 are constraint to be invariant, while error variance of CS9 is freed.

Nederlandse Samenvatting

Introductie

Dit proefschrift gaat over matchingsprocedures bij Nederlandse universiteiten in relatie tot inschrijvingsgedrag en studiesucces van eerstejaarsstudenten. De centrale vraag is of en in welke mate verschillende typen matchingsprocedures aan Nederlandse universiteiten effectief zijn. In dit proefschrift wordt de effectiviteit van matchingsprocedures gedefinieerd door te kijken naar het doel van matching: de juiste student op de juiste plaats krijgen. Een bijkomend doel van matching is het vergroten van het studiesucces, met name het verkleinen van uitval in het eerste jaar en de tijd die nodig is om een diploma te behalen. Aan de hand van deze doelen wordt een effectieve matchingsprocedure geoperationaliseerd als een procedure die 1) door studenten nuttig wordt geacht bij hun uiteindelijke studiekeuze, 2) studenten die geacht worden risico te lopen op uitval, doet nadenken over het afronden van hun inschrijving, en 3) samenhangt met het studiesucces van eerstejaars studenten.

In Nederland is een verkeerde studiekeuze het afgelopen decennium de belangrijkste reden voor uitval zoals studenten zelf rapporteren (Van den Broek et al., 2020). Studenten die niet passen bij de opleiding hebben een grotere kans op uitval dan studenten die wel een aansluiting met de opleiding ervaren (Feldman et al., 1999). Als men de uitval in Nederland wil terugdringen, moet de nadruk dus liggen op het voorkomen van een verkeerde opleidingskeuze en daarom hebben universiteiten hun matchingsprocedures ontworpen met dit doel in het achterhoofd. In dit proefschrift betogen we dat matchingsprocedures waarschijnlijk een “goede” opleidingskeuze bevorderen als ze studenten in staat stellen te testen of ze bij de opleiding passen.

Het onderzoek in dit proefschrift gebruikt de *persoon-omgevingsfit* theorie als leidraad (Lewin, 1935). *Persoon-omgevingsfit* wordt gedefinieerd als de congruentie tussen individuele en omgevingskenmerken (Kristof-Brown & Guay, 2011). Onderzoek waarin wordt gekeken in hoeverre een persoon en diens omgeving bij elkaar passen in verschillende domeinen toont aan dat de prestaties van een individu verbeteren als er congruentie is tussen een persoon en diens omgeving (Ward & Brennan, 2020). Binnen de onderwijscontext bouwt *persoon-omgevingsfit* voort op de veronderstelling dat studenten met bepaalde kenmerken eerder voor bepaalde opleidingen zullen kiezen (Astin, 1993) en dat congruentie tussen student en opleiding van groot belang is voor studiesucces (Feldman et al., 1999). Wij veronderstellen dat studenten die kunnen testen of ze bij de opleiding van hun keuze passen voor ze zich inschrijven, een betere keuze zullen maken.

Welke aspecten van *fit* belangrijk zijn voor een keuze in het hoger onderwijs is niet duidelijk en er is weinig onderzoek over dit onderwerp. Voor dit proefschrift integreren en bouwen we voort op verschillende theorieën, waarvan twee motivatietheorieën, *Expectancy-Value Theory* (EVT; Eccles & Wigfield, 2002) en *Self-Determination Theory* (SDT; Deci & Ryan, 1985), en Tinto's *Student Integration Model* (1977) het belangrijkst zijn. De concepten die we belangrijk achten voor studenten om te bepalen of ze passen in de opleiding van hun keuze zijn 1) zelfperceptie van competentie, 2) interesses, en 3) een gevoel van verbondenheid.

Onderzoeksmethoden

De vier in dit proefschrift gepresenteerde studies zijn gebaseerd op gegevens van verschillende Nederlandse universiteiten. Aangezien informatie over matchingsprocedures alleen beschikbaar is binnen de instellingen zelf, was de bijdrage van deze universiteiten van cruciaal belang voor de vergelijking van de verschillende typen matchingsprocedures die in dit proefschrift worden beschreven. Vier universiteiten stemden ermee in om deel te nemen aan zowel de kwalitatieve als de kwantitatieve onderzoekcomponenten van dit proefschrift. Binnen deze universiteiten kozen we vier opleidingen die tot op zekere hoogte representatief waren voor a) een typische alfa opleiding, b) een typische bèta opleiding, c) een typische gamma opleiding, en d) een opleiding die vaker dan gebruikelijk wordt gekozen door studenten die niet weten wat ze willen studeren.

Deze opleidingen boden een verscheidenheid aan matchingsprocedures aan, waardoor vergelijkingen tussen typen matchingsprocedures mogelijk waren. Het doel was verschillende typen procedures te vergelijken, maar er is een sterke overlap tussen de typen matchingsprocedures en de universiteiten die deze typen aanbieden. In de praktijk betekent dit dus dat een vergelijking van typen matchingsprocedures soms ook een vergelijking van universiteiten inhoudt. Om de anonimiteit te waarborgen, zullen universiteiten en opleidingen in de afzonderlijke studies in dit proefschrift worden aangeduid met nummers en/of letters.

Resultaten

In dit proefschrift onderzochten we de effectiviteit van verschillende typen matchingsprocedures zoals ervaren door studenten (Hoofdstuk 2) en in relatie tot definitieve inschrijving (Hoofdstuk 3) en eerstejaars studiesucces (Hoofdstuk 4 en Hoofdstuk 5).

In Hoofdstuk 2 onderzochten we de perceptie van studenten met betrekking tot verschillende elementen van de Nederlandse matchingsprocedures. Matchingsprocedures dienen als een laatste check op de studiekeuze van studenten. We interviewden 61 aankomende studenten van verschillende opleidingen aan vier universiteiten over hun perceptie met betrekking tot de elementen van matching (d.w.z. vragenlijst; activiteit in de vorm van een persoonlijk gesprek, online studiemodule, of matchingsdag; feedback of advies) in relatie tot het doorzetten van hun voorlopige studiekeuze. Uit de resultaten bleek dat studenten denken dat deze matchingsprocedures een bijdrage kunnen leveren aan het inzicht dat studenten in de opleiding krijgen, en aan het bevestigen van hun studiekeuze. De mate waarin matching kan bijdragen aan deze aspecten verschilt echter tussen elementen van de matchingsprocedures, en tussen groepen studenten. Interessant is dat de twee elementen die deel uitmaken van de matchingsprocedure bij (bijna) elke universitaire opleiding, de vragenlijst en het advies, het minst nuttig werden geacht door toekomstige studenten bij hun studiekeuze. Matchingsactiviteiten (persoonlijk gesprek, online studiemodule, of matchingsdag) werden nuttiger geacht bij het maken van een definitieve studiekeuze dan de vragenlijst en het advies, maar er waren verschillen tussen de verschillende soorten activiteiten. Over het algemeen geldt dat hoe meer aspecten van de *persoon-omgevingsfit* een student kon testen, hoe nuttiger de activiteit werd gevonden.

S

In Hoofdstuk 3 bestudeerden we de relatie tussen typen matchingsprocedures en het percentage studenten dat hun vooraanmelding omzette in een definitieve inschrijving van dertien opleidingen bij vier universiteiten. We beargumenteerden dat hoe lager de inschrijvingspercentages zijn, hoe meer studenten van gedachten zijn veranderd na deelname aan een matchingsprocedure. De bevindingen van dit onderzoek tonen aan dat naarmate matchingsprocedures intensiever zijn, waarbij meer aspecten van de *persoon-omgevingsfit* kunnen worden getest, inschrijvingspercentages lager zijn. Wij toonden aan dat de invoering van de matchingsprocedures in het algemeen, en de intensiteit van deze procedures in het bijzonder, een waarschijnlijke verklaring is voor de lagere inschrijvingspercentages, omdat de inschrijvingspercentages sinds de invoering van de matching zijn gedaald voor opleidingen met intensieve procedures, maar niet voor opleidingen die alleen vervolgactiviteiten aanbieden voor studenten die worden aangemerkt als risicostudenten op basis van hun antwoorden op de vragenlijst.

In Hoofdstuk 4 onderzochten we de relatie tussen de indicatoren van *fit*, zoals gemeten vóór de inschrijving in de matchingsvragenlijsten, en het eerstejaars studiesucces voor alle matchingsopleidingen van drie cohorten studenten aan twee universiteiten. Bovendien onderzochten we of de voorspellers van studiesucces verschillen tussen verschillende studierichtingen. Voor dit onderzoek schatten we eerst een structureel vergelijkingsmodel (SEM) op de gegevens van één universiteit en repliceerden we de bevindingen vervolgens met gegevens van een tweede universiteit. We toonden aan dat indicatoren van *fit*, gemeten voorafgaand aan de start van de studie, voorspellend zijn voor het behaalde eerstejaars gemiddelde cijfer en het aantal behaalde studiepunten. Hoewel de effectgroottes klein zijn, vonden we bijna identieke patronen bij de twee universiteiten in deze studie. Dit sterkt ons in onze overtuiging dat studiesucces in het eerste jaar kan worden voorspeld met behulp van indicatoren van *fit*, gemeten voorafgaand aan de inschrijving. Een tweede belangrijke bevinding is dat de meeste indicatoren van *fit* in ons onderzoek in sterkte en/of richting verschilden tussen bèta- en alfa/gammaopleidingen. Voor zowel bètastudenten als alfa/gammastudenten is het gemiddelde cijfer op de middelbare school de sterkste voorspeller voor studiesucces in het eerste studiejaar. Echter, onze bevindingen geven aanwijzingen dat niet-cognitieve indicatoren, zoals consciëntieusheid en interesse in de opleiding, sterkere voorspellers zijn van studiesucces in het eerste jaar voor alfa/gammastudenten dan voor bètastudenten.

In Hoofdstuk 5 onderzochten we voor motivatieteksten uit matchingsvragenlijsten van alle gestarte studenten aan machtingsopleidingen voor twee cohorten aan één universiteit of tekstmining technieken gebruikt kunnen jaar worden om uitval in het eerste te voorspellen. De resultaten lieten zien dat het antwoord op die vraag tweeledig is. Motivatieteksten in de matchingsvragenlijsten voorspelden eerstejaars uitval even goed als een set van studentenkenmerken waarin de *fit* indicatoren uit Hoofdstuk 4 zijn opgenomen. Echter, wanneer de motivatieteksten en de studentenkenmerken werden gecombineerd, verbeterde de voorspelling van uitval niet. Aan de ene kant lijkt het gebruik van tekstmining technieken voor de voorspelling van uitval dus veelbelovend. Aan de andere kant zou het feit dat het combineren van de tekst- en numerieke gegevens de voorspelling van uitval in deze studie niet verbetert, erop kunnen wijzen dat zij dezelfde onderliggende concepten meten.

Aanbevelingen voor de Onderwijspraktijk

De resultaten die in dit proefschrift zijn gepresenteerd dragen bij aan kennis over wat werkt bij het verbeteren van de studiekeuze van studenten. Deze resultaten hebben praktische implicaties voor procedures in het hoger onderwijs. Hieronder worden drie suggesties gedaan voor beleidsmakers, universiteitsbesturen, personeel en onderzoekers die betrokken zijn bij (de overgang naar) het hoger onderwijs in Nederland en landen met vergelijkbare onderwijssystemen. De suggesties hebben betrekking op de drie hoofdelementen van de huidige matchingsprocedures in Nederlandse universiteiten: de matchingsvragenlijsten, de verschillende matchingsactiviteiten, en het advies.

Ten eerste is het aan te raden dat de matchingsvragenlijsten alleen op empirie gebaseerde voorspellers van studiesucces bevatten. In Hoofdstuk 4 hebben we laten zien dat prestaties op de middelbare school, consciëntieusheid, interesse in de opleiding, de mate van oriëntatie op de opleiding, leeftijd en geslacht voorspellend zijn voor het eerstejaars gemiddelde cijfer en behaalde studiepunten. Bovendien hebben we in Hoofdstuk 5 laten zien dat een schriftelijke motivatietekst geanalyseerd met behulp van tekstmining technieken even voorspellend is voor uitval in het eerste jaar als de combinatie van bovengenoemde kenmerken, maar dat ze elkaar wel lijken aan te vullen op basis van welke criteria worden gebruikt om het model te beoordelen. Daarom zouden deze kenmerken en een schriftelijke motivatie gecombineerd een goede start kunnen vormen voor de matchingsvragenlijsten. Verder kunnen matchingsvragenlijsten afhankelijk van het vakgebied worden aangepast of zodanig worden ontworpen dat ze een algemeen en een opleiding specifiek deel bevatten. Op basis van de bevindingen in ons onderzoek in Hoofdstuk 4 blijkt dat niet-cognitieve voorspellers belangrijk zijn bij de voorspelling van het eerstejaars studiesucces van alfa/gammastudenten, terwijl voor bètastudenten alleen de schoolcijfers in het voortgezet onderwijs de belangrijkste voorspeller zijn. Uit ander onderzoek blijkt dat er mogelijk aanvullende indicatoren voorspellend zijn voor studiesucces, die in deze vragenlijsten kunnen worden opgenomen (zie bijvoorbeeld Van Herpen et al., 2019 die aantoonden dat inspanning in het voortgezet onderwijs een voorspeller is van academisch succes). Het is echter aan te raden om de vragenlijsten kort te houden, omdat aankomende studenten ze niet nuttig vinden bij hun opleidingskeuze. Daarom zijn ze vooral een middel voor het opleidingspersoneel om het risico op uitval van een student in te schatten. Als het opleidingspersoneel de antwoorden van studenten op deze vragenlijsten gebruikt in haar matchingadvies en aankomende studenten daarnaar handelen, kunnen de vragenlijsten bijdragen aan het eerstejaarsstudiesucces in het Nederlandse hoger onderwijs. Echter, als men wil bereiken dat studenten het matchingsadvies serieus in overweging nemen, zijn er wellicht ook enkele aanpassingen nodig in de manier van communiceren van het matchingsadvies (zie hieronder).

Ten tweede moeten matchingsactiviteiten zo worden opgezet dat het mogelijk is voor studenten om te testen of ze passen bij de opleiding. Dit heeft verschillende concrete implicaties. Ten eerste moet een matchingsprocedure uit meer bestaan dan alleen een matchingsvragenlijst en advies. Matching is in de eerste plaats bedoeld om studenten te helpen bij hun opleidingskeuze en zowel de matchingsvragenlijst als het advies worden in dit proces door aankomende studenten niet als nuttig ervaren. Hoe meer aspecten van *persoon-omgevingsfit* getoetst kunnen worden in een matchingsprocedure, hoe nuttiger aankomend studenten deze vinden. In ons onderzoek waren dat online studiemodules en, vooral, matchingsdagen. Gezien het belang van het ervaren je bij de opleiding thuis te voelen, zouden opleidingen die online matchingsprocedures aanbieden moeten overwegen of ze enige vorm van contact tussen aankomende studenten in hun procedures kunnen opnemen, bijvoorbeeld in de vorm van het opzetten van een soort chatroom of een vraag-en-antwoordsessie met huidige studenten. Een derde matchingsactiviteit die we in dit proefschrift hebben bestudeerd, betrof persoonlijke interviews. Hoewel aankomende studenten interviews als een relatief nuttig middel ervaren in hun opleidingskeuze, zijn ze tijdsintensief en dus duur. Bovendien heeft eerder onderzoek aangetoond dat interviews slechte instrumenten zijn voor het voorspellen van het eerstejaars gemiddelde cijfer (e.g., Dana et al., 2013; Reumer & Van der Wende, 2010). Gezien hun gebrek aan voorspellende kracht, kunnen tijd en geld besteed aan het houden van persoonlijke interviews waarschijnlijk beter op een andere manier worden besteed.

Ten derde kan het systeem van het adviseren van studenten worden heroverwogen. Hoewel er een duidelijk verband is tussen een negatief advies en uitval van studenten, negeren veel studenten het negatieve advies (Warps et al., 2017). Dit is vooral het geval, als aankomende studenten het gevoel hebben dat de matching niet representatief is voor de opleiding (Soppe et al., 2019, Hoofdstuk 2). Over het algemeen geven medewerkers nauwelijks negatieve adviezen en ook een advies waarin twijfel wordt geuit, komt niet vaak voor. Leenheer (2022) formuleert dit, vertaalt vanuit het Engels, als volgt: *"De matchingscoördinator is er niet op uit om heel streng te zijn, maar om alleen die studenten eruit te filteren bij wie ernstige twijfels bestaan"*. Medewerkers willen met dit advies een signaal afgeven, maar juist studenten die een niet-positief advies krijgen, vinden dit vaak onterecht en niet nuttig (Warps et al., 2017).

Het zou waardevol zijn om te onderzoeken of de advisering van studenten tijdens matching kan worden verbeterd. Het systeem van concrete advisering (stoplichtanalogieën of positief-negatief systemen) kan worden losgelaten ten gunste van een systeem van specifieker en gedetailleerder feedback geven. De Universiteit Utrecht geeft bijvoorbeeld geen advies, maar generieke feedback waarmee zij hoopt aan te zetten tot reflectie over de opleidingskeuze. Aangezien de matchingsprocedures aan de Universiteit Utrecht als effectief worden beschouwd met betrekking tot de drie criteria voor het evalueren van effectiviteit die in dit proefschrift worden gebruikt, is adviseren van studenten wellicht niet noodzakelijk voor een succesvolle matchingsprocedure. Er kan worden gesteld dat er op het gebied van de advisering van studenten mogelijkheden zijn om de matchingsprocedures verder te professionaliseren en in te passen in de studiebegeleidingssystemen die op de meeste universiteiten al worden toegepast. Als een aankomende student de ontvangen feedback kan bespreken met zijn toekomstige tutor, kan de tutor studenten helpen bij het reflecteren op hun studiekeuze en zullen tutoren direct weten welke studenten misschien extra begeleiding nodig hebben in de overgangsfase naar het hoger onderwijs.

Conclusie

Een van de grootste uitdagingen in onderzoek in het hoger onderwijs is het terugdringen van studie-uitval. In dit proefschrift hebben wij geprobeerd te verklaren hoe matchingsprocedures, die geïmplementeerd zijn om te helpen bij het maken van een studiekeuze en bij het verminderen van uitval, kunnen bijdragen aan dat doel. Een belangrijk resultaat is dat aankomende studenten deze matchingsprocedures nuttiger vinden als meer elementen van de *persoon-omgevingsfit* (d.w.z. zelfperceptie van competentie, interesses, en een gevoel van verbondenheid) kunnen worden getest. Bovendien is het belangrijk dat deze procedures studenten een realistisch beeld geven van de opleiding van hun keuze. Bij opleidingen met matchingsprocedures waarbij meer elementen van *fit* kunnen worden getest, zetten minder studenten hun vooraanmelding om in een definitieve inschrijving. Dit kan betekenen dat hoe meer een matchingsprocedure aan studenten de mogelijkheid biedt om de te testen of ze bij de opleiding passen, hoe meer studenten ervoor zullen kiezen om zich niet in te schrijven wanneer zij een gebrek aan *fit* ervaren. Verder hebben we aangetoond dat indicatoren van *fit*, gemeten in de matchingsvragenlijsten voorafgaand aan de start van de studie, samenhangen met studiesucces in het eerste jaar en dat bepaalde indicatoren (bv. niet-cognitieve eigenschappen zoals consciëntieusheid) naast middelbare schoolcijfers belangrijke voorspellers zijn voor studiesucces van alfa/gammastudenten, terwijl alleen middelbare schoolcijfers voor bètastudenten de belangrijkste voorspellers zijn van studiesucces. Ten slotte lijkt het veelbelovend om tekstmining technieken te gebruiken voor het voorspellen van uitval op basis van motivatieteksten.

About the Author

Karlijn Soppe was born on January 7th, 1992, in Hardenberg, the Netherlands. She obtained her bachelor's degree in Sociology from Utrecht University in 2014 and completed the research master Sociology and Social Research at that same university in 2016.

In August 2016 she started as a PhD candidate and lecturer at the department of Methodology and Statistics of Utrecht University. During her PhD, Karlijn taught in a variety of methods and statistics courses for bachelor students at the Faculty of Social and Behavioral Sciences, including both qualitative and quantitative methods, as well as survey design. In 2021 she obtained the Basic Teaching Qualification for higher education at Utrecht University. In 2019 Karlijn put her PhD research on hold to work as a scientific researcher at the Dutch Inspectorate of Education at the department of Higher Education. For the divisions Selection & Admission and Internationalization she performed several quantitative and qualitative research tasks and co-authored a report on the internationalization of the Dutch Higher Education system. During 2019 and 2020, Karlijn was active as a PhD representative in the PhD council of the Faculty of Social and Behavioral Sciences at Utrecht University.

As of January 2022, Karlijn is working as a postdoctoral researcher at the Centre for Higher Education Governance Ghent (CHEGG) of Ghent University in Belgium.

ICO Dissertation Series

In the ICO Dissertation Series dissertations are published of graduate students from faculties and institutes on educational research within the ICO Partner Universities: Eindhoven University of Technology, Leiden University, Maastricht University, Open University of the Netherlands, Radboud University Nijmegen, University of Amsterdam, University of Antwerp, University of Ghent, KU Leuven, Université Catholique de Louvain, University of Groningen, University of Twente, Utrecht University, Vrije Universiteit Amsterdam, and Wageningen University, and formerly Tilburg University (until 2002).

List update February 10, 2022 (the list will be updated every year in January)

Li, N (08-01-2020) *Analyzing online in-service teacher training courses in China.* Eindhoven: Eindhoven University of Technology

Figueroa Esquivel, F. (13-02-2020) *Early childhood multidimensional development. A rapid and non-linear roller coaster*. Groningen: University of Groningen

De Leeuw, R.R. (27-02-2020) *Through the eyes of the beholder. Unfolding social participation "from within' the classroom*. Groningen: University of Groningen

Schreurs, S. (20-03-20) *Selection for medical school. The quest for validity*. Maastricht: Maastricht University

Nugteren, M.L. (12-06-2020) *What do I know and where do I go? The effects of guidance on task selection*. Utrecht: Utrecht University

Wijns, N. (01-07-2020) *On the hunt for regularities: An investigation of children's early patterning competencies*. Leuven: KU Leuven

Duijzer, A.C.G. (25-08-2020) *Reasoning about graphs in primary mathematics education*. Utrecht: Utrecht University

Dijkema, S. (25-08-2020) R*eady for takeoff? The relation between the type of teacher training program and daily teaching practices of Dutch beginning primary school teachers*. Groningen: University of Groningen

Slot, E.M. (11-09-2020) *Characterizing adolescents' interest: understanding multiplicity and dynamics in persons, objects and contexts.* Utrecht: Utrecht University

Ping, C. (24-09-2020) *Understanding Teacher Educators' Professional Learning* . Eindhoven: Eindhoven University of Technology

Van Rijswijk, M.M. (16-10-2020) *Experiences of continuity and discontinuity in student teachers' development.* Utrecht: Utrecht University

Rovers, S.F.E. (16-09-2020) *Growing knowledge; Supporting students' self-regulation in problem-based learning Maastricht: Maastricht University*

Langeloo, A. (17-09-2020) *Multilingual and monolingual children in kindergarten classrooms: Exploring teacher-child interactions and engagement as learning opportunities*. Groningen: University of Groningen

Sekeris, E. (06-10-2020) *Unravelling computational estimation development in 5- to 7-year-olds* Leuven: KU Leuven

Ter Beek, M. (29-10-2020) *Supporting reading comprehension in history education: The use and usefulness of a digital learning environment* Groningen: University of Groningen

Peppen, L.M. (25-09-2020) *Fostering Critical Thinking. Generative processing strategies to learn to avoid bias in reasoning* Rotterdam: Erasmus University Rotterdam

Van Geel, K., (05-11-2020) *Lifelong learning in radiology: All eyes on visual expertise Maastricht: Maastricht University*

Donker, M.H. (20-11-2020) *In DEPTh: Dynamics of Emotional Processes in Teachers – An exploration of teachers' interpersonal behavior and physiological responses* Utrecht: Utrecht University

Janssen, E.M. (13-11-2020) *Teaching Critical Thinking in Higher Education: Avoiding, Detecting, and Explaining Bias in Reasoning* Utrecht: Utrecht University

Van den Broek, E.W.R. (09-10-2020) *Language awareness in foreign language education. Exploring teachers' beliefs, practices and change processes* Nijmegen: Radboud University Nijmegen

Kasch, J.A. (09-10-2020) *Scaling the unscalable? Interaction and support in open online education* Heerlen: Open University of the Netherlands

Otten, M. (30-10-2020) *Algebraic reasoning in primary school: A balancing act* Utrecht: Utrecht University

De Vrind, E. (25-11-2020) *The SpeakTeach method, Towards self-regulated learning of speaking skills in foreign languages in secondary schools: an adaptive and practical approach* Leiden: Leiden University

Tacoma, S.G. (15-11-2020) *Automated intelligent feedback in university statistics education* Utrecht: Utrecht University

Boonk, L.M. (04-12-20202) *Exploring, measuring, and evaluating parental involvement in vocational education and training* Heerlen: Open University of the Netherlands

Kickert, R. (04-12-2020) *Raising the bar: Higher education students' sensitivity to the assessment policy* Rotterdam: Erasmus University Rotterdam

Van der Wal, N.J. (09-12-2020) *Developing Techno-mathematical Literacies in higher technical professional education* Utrecht: Utrecht University

Vaessen, B.E. (08-01-2021) *Students' perceptions of assessment and student learning in higher education courses.* Eindhoven: Eindhoven University of Technology

Maureen, I.Y. (15-01-2021) *Story time in early childhood education: designing storytelling activities to enhance (digital) literacy development*. Enschede: University of Twente

Van Alten, D.C.D. (19-03-2021) *Flipped learning in secondary education history classrooms: what are the effects and what is the role of self-regulated learning*. Utrecht: Utrecht University

Gestsdóttir, S.M. (08-03-2021) *Observing history teaching: historical thinking and reasoning in the upper secondary classroom.* Amsterdam: University of Amsterdam

Chim, H.Q. (30-03-2021) *Physical Activity Behavior and Learning in Higher Education.* Maastricht: Maastricht University

Krijnen, E. (15-04-2021)*Family literacy in context: Exploring the compatibility of a family literacy program with children's homes and schools.* Rotterdam Erasmus University Rotterdam

Stolte, M. (07-05-2021) *(In)attention for creativity: Unraveling the neural and cognitive aspects of (mathematical) creativity in children.* Utrecht: Utrecht University

Rathé, S. (12-05-2021) *Focusing on numbers – An investigation of the role of children's spontaneous focusing on Arabic number symbols in early mathematical development*. Leuven: KU Leuven

Theelen, H. (12-05-2021) *Looking around in the classroom. Developing preservice teachers' interpersonal competence with classroom simulations.* Wageningen: Wageningen University

De Jong, L.A.H. (20-05-2021*) Teacher professional learning and collaboration in secondary schools.* Leiden: Leiden University

Sincer, I. (20-05-2021) *Diverse Schools, Diverse Citizens? Teaching and learning citizenship in schools with varying student populations.* Rotterdam: Erasmus University Rotterdam

Slijkhuis, E.G.J. (20-05-2021) *Fostering active citizenship in young adulthood.* Groningen: University of Groningen

Groothuijsen-Vrancken, S.E.A. (02-06-2021) *Quality and impact of practice-oriented educational research*. Utrecht: Utrecht University

Hingstman, M. (07-06-2021) *Supporting struggling students: prevention and early intervention with Success for All*. Groningen: University of Groningen

Gerdes, J. (14-06-2021) *All inclusive? Collaboration between teachers, parents and child support workers for inclusive education in prevocational schools.* Amsterdam: Vrije Universiteit Amsterdam

Bai, H. (18-06-2021) *Divergent thinking in young children*. Utrecht: Utrecht University

Wijnker, W. (23-06-2021) *The Unseen Potential of Film for Learning: Film's Interest Raising Mechanisms Explained in Secondary Science and Math*. Utrecht: Utrecht University

Brummer, L. (24-09-2021). *Unrooting the illusion of one-size-fits-all feedback in digital learning environments.* Groningen: University of Groningen

Veldman, M.A. (01-07-21) *Better together, social outcomes of cooperative learning in the first grades of primary education*. Groningen: University of Groningen

Wang, J. (06-07-2021) *Technology integration in education: Policy plans, teacher practices, and student outcomes.* Leiden: Leiden University

Zhang, X. (06-07-2021) *Teachers' teaching and learning motivation in China*. Leiden: Leiden University

Poort, I.C. (02-09-2021) *Prepared to engage? Factors that promote university students' engagement in intercultural group work*. Groningen: University of Groningen

Guo, P. (07-09-2021) *Online project-based higher education Student collaboration and outcomes.* Leiden: Leiden University

Jin, X. (21-09-2021) *Peer feedback in teacher professional development.* Leiden: Leiden University

Atherley, E.N. (27-09-2021) *Beyond the struggles: Using social-developmental lenses on the transition to clinical training.* Maastricht: Maastricht University

Martens, S.E. (15-10-2021) *Building student-staff partnerships in higher education.* Maastricht: Maastricht University

Ovbiagbonhia, R. (08-11-2021) *Learning to innovate: how to foster innovation competence in students of Built Environment at universities of applied science*. Wageningen: Wageningen University

Van den Boom-Muilenburg , S.N. (11-11-2021) *The role of school leadership in schools that sustainably work on school improvement with professional learning communities*. Enschede: University of Twente

Sachishal, M.S.M. (11-11-2021) *Science interest - Conceptualizing the construct and testing its predictive effects on current and future behavior.* Amsterdam: University of Amsterdam

Meeuwissen, S.N.E. (12-11-2021*) Team learning at work: Getting the best out of interdisciplinary teacher teams and leaders*. Maastricht: Maastricht University

Keijzer-Groot, A.F.J.M. (18-11-2021) *Vocational identity of at-risk youth – Tailoring to support career chances*. Leiden: Leiden University

Wolthuis, F. (25-11-2021) *Professional development in practice. Exploring how lesson study unfolds in schools through the lens of organizational routines*. Groningen: University of Groningen

Akkermans-Rutgers, M. (06-12-2021) *Raising readers. Stimulating home-based parental literacy involvement in the first, second, and third grade*. Groningen: University of Groningen

Hui, L. (06-12-2021) *Fostering Self-Regulated Learning: The Role of Perceived Mental Effort*. Maastricht: Maastricht university

Jansen, D. (08-12-2021) *Shadow education in the Netherlands: The position of shadow education in the educational landscape and students' school careers*. Amsterdam: University of Amsterdam

Kamphorst, F. (15-12-2021) *Introducing Special Relativity in Secondary Education*. Utrecht: Utrecht University

Eshuis, E.H. (17-12-2021*) Powering Up Collaboration and Knowledge Monitoring: Reflection-Based Support for 21st-Century Skills in Secondary Vocational Technical Education*. Enschede: University of Twente

Dankwoord

Suavis laborum est praeteritium memoria[16]
CICERO

Eindelijk, nu zijn we klaar. Alle studies bij elkaar. Alleen mijn ouders zullen de verwijzing naar mijn favoriete kinderboek herkennen, maar vele PhDs die mij zijn voorgegaan, zullen zich kunnen vinden in het gevoel. Mijn boek, mijn onderzoek van de afgelopen jaren, eindelijk is het af. En zover was het niet gekomen zonder de hulp en steun van velen.

Allereerst wil ik mijn promotoren bedanken. Ik voel me enorm bevoorrecht om al deze jaren te zijn begeleid door zo'n powerteam. Leoniek, bedankt voor je uitstekende begeleiding. Ik bewonder je passie voor onderzoek én onderwijs, als ook je onderzoek naar onderwijs. Bovendien waardeer ik het enorm hoe je me hebt gesteund en gestimuleerd in al mijn uitstapjes tijdens mijn PhD, zoals mijn detachering bij de onderwijsinspectie, het volgen van het BKO-traject, en zeker ook het starten van een nieuwe baan voordat ik klaar was met mijn promotieonderzoek.

Irene, hoewel je geen expert bent op het gebied van onderzoek naar onderwijs, heb ik je input altijd enorm gewaardeerd. Jij wist altijd precies aan te wijzen waar de hiaten in mijn verhaal zaten of waar mijn betoog voor de (semi)leek niet meer te volgen was. Uiteraard heb ik ook veel gehad aan je uitgebreide statistische kennis, dank daarvoor.

Theo, ik voel me vereerd dat je me hebt willen begeleiden. Officieel was je al met pensioen toen ik aan mijn PhD begon, maar dat weerhield je er niet van om je aan dit – in eerste instantie 5 jaar durende – traject te verbinden. Jouw schat aan kennis over matching, het hoger onderwijsveld, de tijdschriften en hun lezers, is voor mij van onnoemelijke waarde geweest. Bedankt voor al je input en kennisoverdracht.

D

[16]This quote, often attributed to Cicero, can be translated to “sweet is the memory of past labor (or trouble)”.

Als parttime docent bij M&S heb ik het genoegen gehad om met heel veel verschillende collega's samen te werken in grotere en kleinere cursussen. Jeltje, ik heb onbeschrijfelijk veel respect voor hoe jij cursussen als KOM en TOE georganiseerd laat verlopen. Ik heb het altijd fijn gevonden om met je samen te werken. Marieke, bedankt voor je begeleiding bij het geven van online hoorcolleges in Covid-tijd en bedankt voor de gezellige, gezamenlijke conferentiebezoeken. Antonie, ik vond het heerlijk om docent te mogen zijn binnen jouw cursus (die vaker van naam is dan ik me kan herinneren). Ooit was jij mijn werkgroepdocent voor deze cursus, de laatste jaren vormden wij samen de vaste kern van het docententeam. In deze cursus voelde ik me als Socioloog zijnde als een vis in het water. Dank ook aan alle andere (oud)collega's met wie ik heb samengewerkt, met name wil ik hier Boukje, Corine, Marieke H. Nijs, Nina, Peter L., Sjoerd G. en Vera noemen. Jullie hebben me mede gevormd tot de docent die ik nu ben.

Ook zonder mijn paranimfen had ik dit niet gekund. Roeliene, we kennen elkaar inmiddels alweer 10 jaar en spreken elkaar bijna dagelijks. Toch hebben we altijd wel iets om over te kletsen. Ik hoop dat we onze opstart-routine volhouden tot ik bij jouw verdediging kan zijn. Ik heb alle vertrouwen in je, ga zo door. Anne, mijn kamergenootje vanaf dag één. Bedankt voor al je steun, onze SBS-koffie meetings, maandagmorgen WIDM-bespreking en gezelligheid op kantoor. Ik heb je gemist het laatste jaar.

I would also like to thank my fellow PhDs and junior teachers for all the (pre-Covid) lunches, jenga sessions, or quick coffee's. In het bijzonder natuurlijk Hidde en Lientje, tijdens de lockdown werden jullie mijn nieuwe, virtuele, kamergenootjes. Jullie hebben me door de eerste lockdown gesleept. Hoewel het onmogelijk is om van jullie te winnen met pictionary, geniet ik altijd enorm van onze koffiemomentjes.

Onze afdeling functioneert niet zonder het geweldige team van de support staff. Kevin, jij bent de belangrijkste schakel binnen de afdeling, bedankt voor je gezelligheid tijdens de koffiemomenten en je ondersteuning bij mijn eindeloze contractaanpassingen en -verlengingen. Els, bedankt dat je altijd met me meedacht over mijn onderwijstaakstelling, zeker in mijn laatste jaar vanuit Gent; en bedankt voor de fijne samenwerking omtrent de lunch meetings. Chantal, Flip, Irma en Marianne, tijdens mijn PhD waren jullie samen de vaste gezichten van de afdeling, bedankt voor de allround ondersteuning, de gezellige koffiemomenten en de koekjes. Ik wil hier tot slot ook nog mijn geweldige student-assistenten noemen, die enorme hoeveelheden data voor me hebben verwerkt. Camiel, Marlyne en Anne-Roos, jullie waren geweldig.

Natuurlijk wil ik ook al mijn collega's van de afdeling Hoger Onderwijs van de Inspectie van het Onderwijs bedankten. Speciale dank gaat uit naar mijn collega's van team S&T, onder leiding van Perry; team IDM onder leiding van Susan; Ellen vanuit team internationalisering; Karin als mijn leidinggevende en Bianca & Gonneke voor alle praktische ondersteuning. Ik onthoud me hier van het noemen van namen, om niemand te vergeten. Jullie hebben me namelijk als afdeling, stuk voor stuk, met open armen ontvangen. Dus voor iedereen bij HO, hartelijk dank voor het kijkje in de keuken, ik heb met enorm veel plezier met jullie samengewerkt.

Mijn nieuwe collega's bij CHEGG kan ik hier ook niet ongenoemd laten. Jeroen, bedankt voor de kans die je me gegeven hebt door je op te nemen in je team nog voordat ik klaar was met mijn PhD. Davide, bedankt dat je mijn 'peter' wilde zijn. Ik kijk uit naar een fijne samenwerking met jullie beiden.

Daarnaast, niet te vergeten mijn vrienden. Caroline, Liselotte, Nienke, bedankt voor alle avondjes chocofondue en andere gezelligheden. Bij jullie kan ik altijd even ontstressen. Jeroen en Femke, jullie kennen me door en door. Met jullie is het altijd gezellig en ik hoop dat dat nog lang zo mag blijven. Sabine, zo'n 10 jaar geleden begonnen we samen aan een onderwijsavontuur dat ons op vele vlakken gevormd heeft. Ik vind het altijd fijn om met je te praten over lesgeven en allerlei aanverwante zaken.

Tot slot wil ik natuurlijk mijn familie bedanken. Zonder jullie was deze hele reis onmogelijk geweest. Mam, bedankt voor je steun, inhoudelijke discussies, en het lezen én becommentariëren van al mijn papers. Ik kan niet wachten tot je zelf Dr. bent, ik ben trots op je. Pap, bedankt dat je altijd voor me klaarstaat met raad en daad en voor je nieuwsgierigheid naar de praktische toepasbaarheid van mijn onderzoek in je eigen baan als decaan. Lieke, jij laat me zien dat werk niet het hele leven is. Teun, jij laat me zien hoe je werk wel je leven kán zijn. Lieke, ik bewonder op hoe jij je carrière hebt opgebouwd en je onvermoeibaar inzet voor je cliënten, en daarnaast ook nog tijd overhoudt voor je gezin en vriendinnen. Teun, ik ben trots op het feit dat je je droombaan bij defensie niet opgeeft, ondanks alle fysieke tegenslagen die je hebt gehad. Harmen en Kelly, bedankt voor de gezellige verjaardagen, feestdagen en familieweekenden en natuurlijk de eindeloze hoeveelheid spelletjes. Hidde, ondeugend lachebekje, ik hoop dat je net zo'n warm persoon wordt als je ouders.

Of course I also want to thank my Egyptian family. Hala, Mohammed, thank you for always making me feel at home in Cairo and for the lovely family trips.

عزيزتي فوزية
كلما كنا في مصر، تشعرينني كأنني ملكة
الإفطار في السرير، الوجبات اللذيذة ووجودك المحب بشكل عام تعطيني الشفاء الكامل.
تفهمين أننا نحتاج إلى الراحة من ”جدول العمل الأوربي المجنون“.[17]

Lastly, Waleed. You are my rock, my everything. I would not have made it without you. Whenever I was overwhelmed, you calmed me down. Whenever I worked too hard, you supplied me with coffee, tea, and snacks. You annoy the hell out of me while making me laugh at the same time. I cannot imagine my life without you. انا بحبك اكثر

[17]For the curious one's among you, it translates to something like: Fawzya, whenever we're in Egypt, you make me feel like a queen. The breakfast in bed, the extensive meals and your all-round loving presence is simply healing. You understand that we need to rest from "working like those crazy Europeans".