Constructing SIMON: a tool for evaluating personal interests and capacities to choose a post-secondary major that maximally suits the potential

Lot Fonteyne

Supervisor: Prof. Dr. Filip De Fruyt
Co-supervisor: Prof. Dr. Wouter Duyck

GHENT UNIVERSITY

Constructing SIMON: a tool for evaluating personal interests and capacities to choose a post-secondary major that maximally suits the potential.

DISSERTATION ABSTRACT

Choosing a suitable study program is an arduous process for many prospective students. Despite the bulk of information provided by institutions only 40% of enrolling students in Flanders pass all courses in the first year of higher education. Too many students fail in their first year because they are not 'at place'. These students are in need of valid tools that help them choose a study program that maximally suits their interests and potential. This dissertation is aimed at describing the construction and validation of such an internet-based self-assessment tool, SIMON (Study capacities and Interest MONitor).

An instrument such as SIMON needs to answer the two basic questions that prospective students are faced with when going through their study choice process: "what programs interest me?" and "will I be able to succeed?". Therefore, the construction and validation of a new and context-specific interest tool is discussed that allows (prospective) students to answer the first basic question. The second question (will I be able to succeed?) is addressed by examining the predictive validity of a broad range of variables for tertiary academic achievement. The incremental predictive validity of background factors, cognitive skills and the non-cognitive factors of personality, self-efficacy, motivation, metacognition and test anxiety are examined in a large sample of students. Moreover, the differential predictive validity of these variables is examined across different tertiary education programs. This will allow (prospective) students to evaluate their capacities with reference to specific study programs.

Still, answering these two questions is not necessarily enough to get (prospective) students 'in the right place'. A key matter is whether they are activated by the feedback they receive from such an instrument. Therefore, attention is also devoted to the consequential

validity of SIMON by examining the effect of receiving negative attainability feedback on career goal disengagement.

It is concluded that SIMON can help students during their study choice process. Directions for future research and further development of SIMON are also addressed.

The Department of Information and Communication Technology also advanced the project. A special thanks to Thomas Berton, Phillippe Depoorter and Danny Schellemans.

Each and every faculty was crucial to the development and implementation of SIMON. I thank all deans and educational directors for their support and the ability for me to collect so much data. The efforts of the study guidance counselors across all faculties were also indispensable. Also thanks to Stijn and Elisabeth, who joined the team and are continuing our important mission.

The colleagues at the Department of Educational Support of the Faculty Psychology and Educational Sciences deserve specific attention. Thanks to Daphne, Tom and Kris, the "original gang", but of course to all other FDO-members. I appreciate each and every one of you and I will definitely miss our lunches and other activities!

Outside of Ghent University, I would like to thank the Association partners HoGent, Artevelde University college and Howest; and the Vrije Universiteit Brussel for their efforts for data collection.

The people in my personal environment and their support has also been vital. A special thanks to my mother, to all my brothers, sisters and friends and to my cousins Justine and Valentine.

Wendy, of course, for just about everything! Thank you for your unrelenting support, I really could have not done this without you! Although SIMON was my 'first baby', I am even more proud of our shared project: our twin daughters. Liv and Nell, you are amazing. I hope I can continue to assist you in becoming who you are and what you want to be, whatever that may be!

Lot Fonteyne, Ghent, May 2017

TABLE OF CONTENTS

Chapter 1: Introduction and overview

The overall aim of this dissertation is to document the construction and validation of a tool that provides prospective students with the necessary information to choose a higher education study program that maximally suits their interests and potential.

This introductory chapter describes the research context and more specifically the educational system in Flanders, the northern region of Belgium which has an autonomous educational system. We describe the growing awareness of the necessity to develop an instrument that aids potential students in their choice of higher education study program and allows to understand the context in which the specific components of the tool were developed. In doing so, we elaborate on the study choice process theory and on potential problems associated with this process. We describe how a tool that aids study program choice can counter these issues. Finally, an outline is given of the main components of the tool, of the data collection process and of the specific research hypotheses that are addressed in the present dissertation.

<div align="center">

The Educational System in Flanders

</div>

In order to understand the necessity of constructing a tool that helps prospective students to choose a higher education study program it is important to elaborate on contextual factors. Especially given that the Flemish educational system is quite distinct from systems across the world, more specifically with regards to entrance requirements for higher education. The majority of countries and regions in the world apply some form of selection at the entry to tertiary education (McGrath et al., 2014). Whether it is through standardized aptitude tests (e.g., Japan, Sweden, Turkey, U.S.), centralized secondary school exit exams (e.g., Australia, France, Germany, Italy, U.K.) or through other entry requirements such as grade point average, interviews, portfolios, and application essays (McGrath et al., 2014), most regions apply some form of selection of higher education students. In contrast, the Flemish higher education system

is almost maximally unconstrained. With the exception of medical, dentistry and performing arts programs, there are no selection exams or admission tests. The sole requirement for enrollment in any other program is holding a secondary education qualification. And even students without qualification can be granted access. Also, this secondary education qualification is not obtained through any centralized or standardized examination, as is the case in many other open access systems. In Flanders, it is the class committee (consisting of the head teacher and all other teachers who teach the pupil) that decides whether or not the pupil has sufficiently achieved the objectives of the curriculum and thus passes or not (Flemish Ministry of Education and Training, 2008).

As the necessity to develop an orientation tool is a product of the regional context, it is imperative to further delineate the structure of both higher and secondary education in Flanders and the implications of the open access policy for academic achievement.

Secondary education structure

Figure 1 depicts the structure of upper secondary and higher education in Flanders. There are four types of secondary education (SE) programs: general, arts, technical and vocational SE.

General SE has an emphasis on broad general education and provides a solid foundation for higher education (95% of General SE students pass on to higher education; Van Daal, Coertjens, Delvaux, Donche, & Van Petegem, 2013). Technical SE emphasizes general and technical matters and prepares for a profession. Passing on to higher education is possible, but less frequent (69.1%; Van Daal et al., 2013). Secondary arts education combines a broad general education with active arts practice and also prepares for a profession or to pass on to higher education. Finally, vocational SE is a practice-oriented education in which young people learn a specific profession (Flemish Ministry of Education and Training, 2008), after which higher education is less likely (23.6%; Van Daal et al., 2013).

UNIVERSITY COLLEGES
VOCATIONAL TRACK/TERIARY TYPE-B

UNIVERSITIES
ACADEMIC TRACK/TERIARY TYPE-A

AGE

TERTIARY EDUCATION

SECONDARY EDUCATION

Doctor/Phd

Master following master

Bachelor following bachelor

Master

Professional bachelor (vocational track/tertiary type-B)

Bridging program

Academic bachelor (academic track/tertiary type-A)

18

16

Year 6	Year 6	Year 6	Year 7
Year 5	Year 5	Year 5	Year 6
			Year 5
General	Arts	Technical	Vocational

Fig. 1. Upper secondary and tertiary education structure in Flanders, including admission requirements (arrows).

Although the four types of secondary education have different content and emphasis and differ with regards to their finality and the extent to which they prepare either for further education or for the job market, admission to higher education programs is independent of the type of secondary education qualification obtained.

Higher Education Structure

Flemish higher education can be described as binary (Arum, Gamoran, & Shavit, 2007). It consists of two main types of programs: academic and professional/vocational (see Figure 2 for a graphical representation of the higher education structure). Academic programs are mainly organized by universities, whereas university colleges provide professional programs with an emphasis on functional skills. While the focus in the latter is more on concrete and specialized professional skills and direct entry into the labor market, academic programs are more theoretical and research-oriented, leading to a master degree. The professional programs lead to a bachelor degree and correspond to the Bologna first cycle programs of 180 European Credit Transfer and Accumulation System (ECTS) ("The Bologna Declaration of 19 June 1999. Joint

declaration of the European Ministers of Education", 1999). Academic programs also lead to a bachelor degree at first (which also consists of 180 ECTS credits), but the finality is to complement this degree by a master. Academic programs thus correspond to the Bologna two-cycle programs (for a detailed description of the higher education system in Flanders, we also refer to Kelchtermans & Verboven, 2008). These two higher education tracks correspond to the distinction between tertiary-type A (or academic) and tertiary-type B (or professional/vocational) programs as specified in the International Standard Classification of Education (UNESCO, 1997).

Although the academic track is well-represented in Flanders, the vocational track for a first degree is relatively more popular than in most other countries. Across countries, on average 39% of young people will graduate from tertiary-type A first-degree programs (often called bachelor's degree) and 18% from tertiary-type A second degree programs (often called master's degree) (Organisation for Economic Co-operation and Development: OECD, 2014). Compared to these averages, fewer people (only 18%) in Belgium attain a first degree in tertiary-type A education but more people (26%) will graduate from tertiary-type A second degree programs (master's degree). This is possible because the lower tertiary-type A first-degree graduation rate in Belgium is counterbalanced by a higher level of first-degree graduation rates from tertiary-type B (vocationally oriented) programs (32% compared to the OECD-average of 14%). Belgium is one of the only countries (next to Argentina and Slovenia) in which more people earned their first degree from tertiary-type B programs than from tertiary-type A programs (OECD, 2011).

At the end of secondary education (age 17-18), students are expected to decide on which study program they want to pursue. This choice entails both the study level (either academic or vocational track) and the study field or major (e.g., engineering, law, psychology, foreign languages…). With very few exceptions, study fields can be studied either at the theoretical or

at the more applied level. For example, the academic Psychology program extensively studies the fundamental principles underlying human psychology, hereby considering different theoretical perspectives, as well as the development of research competencies relevant for the scientific field, whereas the vocationally oriented 'Applied Psychology' program focuses on the practical application of psychological principles. As said, students in Flanders choose a major when enrolling in higher education. When a student wants to change majors this usually requires him or her to start over and re-enroll as a freshman. This is in contrast with systems that allow undergraduates to take courses across several disciplines before choosing one major field of study in which to specialize (as is typical for instance in the U.S.).

The academic year and evaluations.

Once enrolled, the academic year starts at the end of September and it consists of two semesters. Courses usually take one semester and students are evaluated at the end of each semester during a first exam period (in January for the first semester and June for the second). Many courses, especially in the first bachelor year, are evaluated through written exams with a multiple choice or, less frequently, an open answer format. In about 10 to 20% of the first year courses, these exams are complemented with coursework and participation credits. A student passes the course when he or she earns a score of minimum 10 out of a maximum of 20. When students fail a first time there is a second examination period at the end of the same academic year (in August). If students do not achieve the minimum during this second examination chance, they fail the course.

Academic Achievement in the First Year of Higher Education

On average, Flemish students earn 61% of their ECTS credits in their first year of tertiary education. A mere 40% of students pass all courses in the first year and 17% does not earn one single credit (Ministerie van Onderwijs en Vorming, 2009). Only 38% of the Flemish students who enter a bachelor's program graduate on time. Although this success rate does not

fall far below the average of 41% across OECD countries (OECD, 2016), there is room for improvement as failing a year of higher education carries a high cost. Parents and students not only need to pay the tuition fee and other study-related costs such as transport, housing, food, and study material but they also suffer a loss of income compared to when the student would have entered the labor market. The government and the higher education institution also bear a high financial cost. In Belgium, the 2011 public expenditure per tertiary education student was 11,599€ (EUROSTAT, 2017) and recent OECD-data (2016) shows that a Belgian student costs a higher education institution15,911 USD per year.

But there is also a high personal cost. Students who perform badly have a higher risk of dropping out of tertiary education which in turn has individual, economic and social consequences. People with lower educational attainment generally have worse health, are less socially engaged, have lower life satisfaction, lower employment rates and lower relative earnings (OECD, 2016).

Thus, the cost of failing in higher education is high for parents, students, institutions and the government (Declercq & Verboven, 2010), which makes it very relevant to try to improve success rates.

The Necessity to Develop a Tool

In sum, the organization of education in Flanders guarantees a fairly unrestricted access to higher education. Moreover, there is a policy of high government funding and low tuition fees (Kelchtermans & Verboven, 2008), which are typically below €1000/year. This system is assumed to guarantee socially fair access and to improve participation of economically disadvantaged groups in higher education, but the open entrance implies that the first year of university is typically a "selection year". This is demonstrated by the fact that only 40% of students pass all courses during the first year of studying and is in line with international

findings: graduation rates in open admission systems are typically lower (32% on average in comparison to the international average of 37.13%) (McGrath et al., 2014).

Oppedisano (2009) hypothesized that the combination of open admission policies and low tuition fees invites young people to experiment with academic studies. To discourage this trial and error choice behavior, she proposes to provide students with better information about their prospects for success. This recommendation is acclaimed by many others such as McGrath et al. (2014) and Vossensteyn et al. (2015). They posit that supplying accurate information prior to enrolment improves the ability to select suitable study routes. Moreover, McGrath et al. (2014) suggest that strengthening the pre-university orientation process can increase social equality in higher education. This may well be the case as it are often socially vulnerable groups that lack the information to make a realistic educational program choice or to enroll in tertiary education (Müller, 2014; OECD, 2003).

Study choice process: choice theory

But what type of information do prospective students require? In 1909 already, Parsons (as cited in Brown, 2002, p.5) set forth the three fundamental factors in making a wise vocational choice: (1) a clear understanding of the *self* (abilities, interests, ambitions); (2) a knowledge of the requirements of the *environment* (conditions of success, advantages and disadvantages, prospects); (3) true reasoning on the *relations of these two* groups of facts.

Since then, this idea of *person-environment fit* (Dawis, 2004; Holland, 1985) has been the fundament of career choice theories such as the theory of circumscription (Gottfredson, 1981, 1996) and career construction theory (Savickas, 2006). The underlying rationale is that students who make a realistic choice will perform better. Research indeed suggests that congruence between person and environment is related to higher levels of educational stability, satisfaction, performance, and persistence of higher education students (Feldman, Smart, & Ethington, 1999; Nye, Su, Rounds, & Drasgow, 2012).

As a result, career choice theorists stress that an optimal career choice process is conditional on the exploration of both the self and the environment. The research on the stages of career decision making suggests that individuals should begin with a broad exploration of talents and interests, continuing with the crystallization of a narrower set of specific career options, and culminating in concrete choices about jobs and careers (Feldman & Whitcomb, 2005). Gati and Asher (2001, p. 142) for example, presented a 3-stage model for career decision-making processes which includes: (1) A prescreening of potentially relevant career alternatives, based on the individual's preferences, to locate a manageable set of alternatives that deserve further exploration; (2) In-depth exploration of the promising alternatives (including an examination of the possibility of actualizing them); and (3) Comparison and choice of the most suitable alternative.

The quality of this study choice process is important for subsequent academic outcomes. Germeijs and Verschueren (2007) for example, found that higher levels of self-exploration and in-depth exploration of the environment at the end of secondary education were beneficial for academic adjustment and commitment to the study at the beginning of higher education.

Study choice process: choice reality

Although career theorists agree on this importance of the career-decision making process in general and the exploration of the self and the environment more specifically, findings on how prospective students actually accomplish their choice are discouraging. The 3-stage model above describes the optimal way of making career choices but the reality of how people actually decide is often rather different (Pitz & Harren, 1980).

For example, Wessel, Ryan, and Oswald (2008) found that the perceived and the objective fit between college students and their major bore little relation to one another ($r =$ .05). They hypothesized that this results from the lack of understanding of themselves or their environment (or both) when choosing a college major. Consequently, students may believe their

interests match certain majors, but their perceptions of those majors, or their perceptions of themselves, differ from the actual person and environment. Similarly, Grotevant and Durrett (1980) established that the occupational knowledge of high school students was very limited. They were especially lacking accurate knowledge of the educational requirements of careers they wished to enter, and knowledge of the vocational interests predominantly associated with their occupational choices. More recently and specifically for Flanders, Van Daal et al. (2013, p.54) found that Flemish secondary education students, even barely three months prior to the start of higher education, had only spent a limited amount of time on exploring their own options and on their choice of study.

Thus, it seems especially appropriate to facilitate informed decision processes in prospective students as these can ensure stronger retention and higher graduation rates. This requires valid and context-specific instruments to aid prospective students in making an informed choice. Unfortunately, until the start of the current project no such tools were available in Flanders.

This dissertation is aimed at describing the construction and validation of an internet-based self-assessment tool, SIMON (Study capacities and Interest MONitor), that supports an optimal study choice by generating honest and valid feedback on both personal attributes and the match with educational possibilities in Flanders.

Components and development process of SIMON

When providing information on the match between a person (prospective study) and the environment (study program), two important personal attributes have been identified as important: interests and competencies (skills and abilities). These attributes correspond to the two main questions young people ask themselves when going through the arduous process of selecting a suitable study program: (1) "what do I want to study?" and (2) "will I be able to succeed?".

The first question concerns the fit between interests and study programs. The main goal in the provision of information on interests-environment fit is to encourage maximal exploration of (relevant) study options. Previous research has demonstrated that student decision makers typically pick initial, intuitively derived choices, and then fail to give serious consideration to other options later in the process (Feldman & Whitcomb, 2005; Krieshok, Black, & McKay, 2009). Therefore, by giving prospective students a list of matching programs based on their personal interests, it is our ambition to broaden their perception of viable options.

The second question pertains to the fit between personal skills and abilities and the study program environment. Feldman and Whitcomb (2005) found that the use of information on the match between abilities and the environment was effective in reducing the set of feasible career alternatives. Thus, whereas the interests component intends to broaden the choice options, the assessment of competencies is aimed at narrowing them down.

The development of SIMON is centered around these two components. Following the particularities of the Flemish educational system (as described above), both components are tailored to this specific context.

Interests

Contents.

The first component of this dissertation pertains to the students' interests and the extent to which these are aligned with particular study programs. Up till now, there is a lack of valid instruments that link students' interests to the available higher education programs in Flanders. Therefore, a first important focus is the development of a context-specific interest assessment tool and feedback module (the SIMON Interest inventory: SIMON-I). Because of the comparably high prevalence of tertiary-type B enrollment in Flanders, this context-specificity especially implies the incorporation of a means to discriminate between interests in the vocational versus interests in the academic track.

10

Conceptual framework.

In designing SIMON-I, we used Holland's (1997) RIASEC interest model as taxonomic framework, which is the most influential model of vocational choice making (Brown, 2002, p.6). Central in Holland's theory is the assumption that both people and environments can be described in terms of their similarity with six different personality and environment types, i.e., Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (for a description of these types, see Nye et al., 2012). The theory postulates that students choose academic environments compatible with their personality types and, in turn, academic environments reward different patterns of student abilities and interests. When applied to study program environments, this implies that Artistic study programs attract and are dominated by Artistic personality types, whereas Social study programs attract Social types. Holland's theory further assumes that satisfaction and achievement of people is a function of the congruence or fit between their personality type and their environment (Feldman et al., 1999).

Following Holland's theory in constructing SIMON-I, it was imperative to characterize both person and environment in terms of RIASEC types. On the person side, this implied the construction of an interest inventory that allows to capture respondents' interests in terms of the underlying RIASEC structure. On the environment side this required the description of all included study programs in terms of RIASEC dimensions, which can be done using different procedures. In the construction of SIMON-I, two main methods were and are being used: the judgment method and the incumbent method (see, Rounds, Smith, Hubert, Lewis, & Rivkin, 1999). The judgment method relies upon the direct rating of occupations by judges or experts whereas the incumbent implies the use of the empirically established scores per program to refine the profiles generated by experts. When applying these procedures, each study program environment receives a RIASEC code which allows for matching between the person (prospective student) and the environment (study programs).

Data collection.

The development process of SIMON-I started in 2012 by constructing a valid inventory to assess the personal interests of prospective students. As we describe in chapter 3, several versions of this inventory preceded the one that is now used. To adequately characterize the environment of study programs, we started off with an expert coding (judgement method) of all programs, but from the very beginning (academic year '12-'13) data collection was initiated to allow the application of more empirical methods such as the incumbent method. This collection encompassed the assessment of interests of successful students across all study programs as we describe in chapter 3. Since then, a new wave of data is collected each year and up till now our dataset consists of 13,535 valid responses across the 5 co-operating institutions, which allows us to refine the study program interest profiles.

Competencies

Contents.

The second component concerns the match between the individual skills and abilities on the one hand and study program requirements on the other. This necessitates assessing relevant personal attributes and linking these to study programs. As SIMON intends to inform potential students on their prospects for success, the focus lies on the predictive validity for academic achievement. In the past, the prediction of academic success has relied heavily on cognitive factors. Still, during the last decades, researchers have evidenced the importance of non-cognitive factors as well (see e.g., Credé & Kuncel, 2008; Lipnevich & Roberts, 2012; Poropat, 2009; Robbins et al., 2004). Therefore, in constructing SIMON we took into account the predictive validity of both cognitive and non-cognitive variables. Because it is likely that different study programs require different (levels of) skills and abilities, we also investigate the importance of making program-specific predictions. Although this SIMON-Competencies part (SIMON-C) also bears resemblance to high-stakes selection and admission tests, it does differ

fundamentally. As opposed to these types of tests that often try to identify excellent students, SIMON intends to identify students who almost certainly lack the necessary skills to pass their first year of higher education. This aligns with the open access policy in Flanders. Consequentially, the focus is on the assessment of very basic abilities and on a high prediction accuracy, especially limiting false negative advice: only a small group of students should get a clear warning that a program is unattainable, but this prediction should be very accurate and it should indicate that a student almost certainly lacks the very basic skills that are necessary to succeed in the first year of higher education. Students who might be able to pass should get the benefit of doubt and should not be discouraged. Moreover, as opposed to high-stakes selection tests, the results of SIMON are not binding. Their primary aim is to raise awareness on the accordance of the individuals' competencies (and interests) with the demands of higher education programs. As such, it aims to support an optimal, but free, choice of study program.

Data collection.

The project started off in the academic year 2011-2012 when the basic mathematics test (described in chapter 3) was first administered in the faculty of Psychology and Educational Sciences of Ghent University in a sample of 502 students. At the end of this year, it became apparent that this basic test was predictive of academic achievement and it was decided to examine whether it could be expanded and transferred to other study programs. First, a thorough review of the literature on academic achievement was undertaken which resulted in the selection of a variety of factors and tests that had shown to be predictive for student success. An overview of the sample size and the included study programs and tests for each cohort is provided in Table 1. These tests were first administered during the academic year 2012-2013 in a sample of students ($N = 532$) restricted to the faculty of Psychology and Educational Sciences. New incoming students were tested at the start of the academic year and their test scores were related to their end-of-year study results with the aim of validating program-

specific predictive models that could be used to advise prospective students. Results of this pilot year were promising and from then on the project developed progressively. From the academic year '13-'14 onwards more and more faculties were engaged to gather data for SIMON-C which implied the inclusion of an increasing amount of study programs and respondents. In response to this expansion, the number of included tests was also raised to incorporate the assessment of more program-specific knowledge on subjects such as chemistry or physics (see Table 1).

In 2015, the board of Ghent University decided to oblige new incoming first year students to fill out SIMON-C. As a result, the response rate in the academic year '15-'16 raised to 81.2% of all incoming students (see Table 1).

Apart from Ghent University, other institutions also collaborated in collecting data for SIMON. A total of 6,045 students completed SIMON-C in Artevelde University College (data collected from '15-'16 onwards), University College Ghent (from '15-'16 onwards), University College West Flanders (from '15-'16 onwards) and Free University Brussels (during the academic year '15-'16). Thus, the SIMON-C component now relies on a sample of 22,008 students across all involved institutions, and data is continuously gathered and used to further perfect, develop and validate the instruments.

Procedure: prediction of academic success.

Validating the prediction of academic success requires the tracking of prospective students from pre-enrollment until they finish their first year of higher education. Ideally, students' skills and abilities would be assessed before enrollment and these would be related to academic achievement after the first year of higher education. Yet, this method poses practical, methodological and legal problems. For example, response rates would probably drop dramatically between pre-enrollment assessment and the end of the first year of higher education. Also, as it is legally very difficult to access study results of students in other higher education institutions, this would force us to work with self-reported achievement measures.

Table 1 Overview of the data collection process for SIMON-C at Ghent University

Academic year	Components	Faculties[1]	N	Response rate
11-12	Basic mathematics test: chapter 3	L; PE; PS	502	90.2
12-13	Previous + Vocabulary knowledge: Lextale[2] + Reading comprehension: SweSAT[3] + Motivation: SRQ[4] + Self-efficacy: CASES[5] + Metacognition: MAI[6] + Test anxiety: CTAS[7]	PE	532	93.4
13-14	Previous + Self-control: SCS[8] + Grit: GRIT-short[9]	AL; L; MH; PE; PS; VM	1351	42.4
14-15	Previous + stronger mathematics test: Newly developed	AL; BE; EA; EB; L; MH; PE; PH; PS; VM	3343	59.4
15-16	Previous + Chemistry: Newly developed + Physics: Newly developed + Conscientiousness: PFPI[10]	All = AL; BE; EA; EB; L; MH; PE; PH; PS; S; VM	5290	81.2
16-17	Previous + Reasoning ability: Newly developed	All = AL; BE; EA; EB; L; MH; PE; PH; PS; S; VM	4945	73
TOTAL			15,963	

[1] AL: Arts and Literature; BE: Bio-engineering; EA: Engineering and Architecture; EB: Economics and Business Administration; L: Law; MH: Medicine and Health Sciences; PE: Psychology and Educational Sciences; PH: Pharmaceutical Sciences; PS: Political and Social Sciences; S: Sciences; VM: Veterinary Medicine

[2] Lemhöfer and Broersma (2012) [3] SweSAT (2011) [4] Vansteenkiste, Sierens, Soenens, Luyckx, and Lens (2009a) [5] Owen and Froman (1988) [6] Schraw and Dennison (1994) [7] Cassady (2004) [8] Tangney, Baumeister, and Boone (2004) [9] Duckworth and Quinn (2009) [10] De Fruyt and Rolland (2010)

To counter these issues, we tracked newly enrolled students by assessing skills and abilities at the very start of the academic year. As such, the responses resemble those of a population of prospective students. This approach allows us to use real study achievement measures, which are collected from the institutional database at the end of the academic year.

To estimate the chance of success in SIMON, we use recursive feature elimination and cross-validation. This procedures are applied for each study program separately, thus generating program-specific chances of success. First, the dataset is split into a 75% training set and a testing set containing 25% of the data. The training set is used for model selection by evaluating the predictive power of explanatory factors for achievement in the first academic year. The testing set is used to measure how well the model performs at making predictions in a different sample. Model selection occurs by applying recursive feature elimination to the training set. Recursive feature elimination is a logistic regression that follows the backward stepwise procedure and is embedded in a K–fold cross–validation. Cross–validation is performed on 10 subsets and is repeated 3 times. This analysis shows how many and what variables should be included in the model. Classification success of the model is usually evaluated using a cut–score of .50. Yet, this does not serve our aim. SIMON intends to classify (prospective) students in three groups, which requires the selection of two different cut-scores. Therefore, we are in search of one cut–score that allows us to identify students at risk of failure, without wrongfully classifying passing students and secondly, we look for another cut-score that identifies students with a high probability of passing. Currently, a sensitivity value of 95% (for the low chance group) and of 70% (for the high chance group) are selected. This means that we allow a fall-out of 5% in the low chance group and of 30% in the high chance group. Thus, the threshold for high probability of passing is more relaxed because students may have all the required prerequisites to pass but still fail because of situational, emotional or behavioral impediments during their first academic year. After the model and the cut-scores are established, this model

is evaluated using cross-validation. Parameter estimates of the logistic regression model are forced onto the testing sample and the diagnostic values of the model are evaluated for the low and the high chance groups. If, again, we find a sensitivity of 95% in the low-chances group and 70% in the high-chances group, the model and the identified cut-scores are retained. Thus, in the application of SIMON, a low chance of passing means that the respondent has a 95% of failing and a high chance of passing indicates a 70% chance of passing. Respondents wo do not fall within these two groups are classified in the 'average' group which means that the prediction of passing is difficult.

Feedback

As stated above, from 2015 onwards the board of Ghent University obliged students to fill-out SIMON-C. This decision was based on the recognition that the data collected for the validation of SIMON could also be beneficial for the identification of newly enrolled students that were at risk of failing their first academic year. Together with this decision, post-enrollment SIMON was born. For the very first time, newly enrolled students who participated in SIMON-C received a personalized feedback report. Thus, although the main target audience of SIMON are potential students on the verge of making a career choice at the end of secondary education, the availability of validation data offers advantages for students who are enrolled too. SIMON may allow students who are already enrolled to get an idea of their starting position in higher education. As such, the SIMON information can also be used to activate enrolled students who have a high likelihood of failing their first year. When identified, these students can get information on remedial activities that might increase their chances of success. If students make use of this information, SIMON can also alleviate student success and retention post-enrollment.

This brings us to a third general component in the current dissertation: giving feedback. Even when the instrument gives feedback on the match of interests and capacities with specific

study programs, the question remains whether (prospective) students are activated by the feedback they receive. If they are not, the instrument does not support the study choice process and is not able to increase higher education success and retention. Validation requires evaluation of how test results are used (Duijnhouwer, Prins, & Stoking, 2012). This consequential validity is an important aspect of construct validity (Messick, 1990). And although this type of validity is indispensable, surprisingly few studies addressed the issue of the action behaviors that result from test reports (Hattie, 2009). The current dissertation also contains an investigation about how SIMON test results are used.

Overview of the Current Dissertation

The sum of the three components (interests, competencies and feedback) leads up to SIMON as an orientation instrument. How these specific topics are examined and implemented is detailed in the following chapters.

In chapter 2, we present an overview of how the research results were implemented as practical tools that aid (prospective) students in their process of choosing a higher education study program. We also elaborate on the technical features of SIMON by providing criterion validity evidence and by examining test fairness issues.

Interests: SIMON-I

In chapter 3, the development, initial validation and practical value of the SIMON Interest tool (SIMON-I) for secondary education students who are in the process of choosing a higher education program is described. SIMON-I is based on John Holland's RIASEC model (Holland, 1997) but also introduces an 'Academic-track scale' which allows to discriminate between interest in academic versus vocational programs across and within fields of study. A sample of 3,962 students is used to evaluate the structural validity of the measure with an additional focus on possible gender differences in item functioning (i.e., differential item functioning) and in structural validity. The criterion validity of the newly proposed Academic-

track scale is addressed and the usefulness and face validity of the SIMON-I output are examined. Special attention is also given to the feedback module of the tool.

Capacities: SIMON-C

In chapter 4, the predictive validity of a test of basic mathematical skills is examined. This newly developed test is easy to administer and is aimed at identifying students who lack the very basic, but necessary skills to successfully take on an introductory statistics course in an academic bachelor program. Because of the heterogeneity of new incoming students and the lack of standardized testing in the Flemish education system, this test can be especially helpful in identifying at-risk students. We examine not only whether this test can predict academic achievement in a statistics course over and above secondary education background, but also whether the test can predict overall first-year achievement.

In chapter 5, a study that assesses the relevance of a broader range of variables is discussed. Instead of focusing solely on mathematical (or cognitive) skills and background factors, the non-cognitive factors of personality, self-efficacy, motivation, metacognition and test anxiety are also taken into account. As such, we evaluate the incremental predictive validity for tertiary academic achievement of this broad range of variables in a large sample of students ($N = 2,391$). Moreover, we examine the differential predictive validity of these variables across different tertiary education programs. If there are disciplinary differences in the predictive power of variables, prospective students would benefit from the opportunity to evaluate their personal skills with reference to specific fields of study as opposed to receiving generalized feedback on their compentence level.

Feedback

In chapter 6, we examine the effect of receiving negative attainability feedback on career goal management. Can negative attainability feedback encourage students to disengage from an unattainable career goal at the start of the university trajectory? How do students react to

negative attainability feedback (as opposed to positive attainability feedback): by doubling their effort (as proposed by control theory) or by exploring other options (as suggested in social cognitive theories)? And are these management strategies mediated by self-efficacy, motivation and the perceived accuracy of feedback? At a more descriptive level, we evaluate to what extent students who receive negative attainability feedback are activated by their feedback report (by putting in more effort for their studies, by participating in guidance activities or by considering to change majors).

Finally, the general conclusions are presented in chapter 7. The research findings are synthesized in light of practical implications. Directions for future research and for further development of SIMON are also addressed.

References

Arum, R., Gamoran, A., & Shavit, Y. (2007). More inclusion than diversion: Expansion, differentiation, and market structure in higher education. In R. A. Y. Shavit, A. Gamoran (Ed.), *Stratification in Higher Education: a comparative study* (pp. 1-35). Palo Alto, CA: Stanford University Press.

The Bologna Declaration of 19 June 1999. Joint declaration of the European Ministers of Education., (1999).

Brown, D. (2002). Introduction to Theories of Career Development and Choice: Origins, Evolution, and Current Efforts. In D. Brown (Ed.), *Career choice and development* (Fourth ed.). San Francisco: John Wiley and Sons, Inc.

Cassady, J. (2004). The influence of cognitive test anxiety across the learning–testing cycle. *Learning and Instruction, 14*(6), 569-592. doi:10.1016/j.learninstruc.2004.09.002

Credé, M., & Kuncel, N. (2008). Study habits, skills, and attitudes. The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science, 3*(6), 425-453. doi:10.1111/j.1745-6924.2008.00089.x

Dawis, R. V. (2004). The Minnesota theory of work adjustment. In S. D. Brown & R. W. Lent (Eds.), *Career development: Putting theory and research into practice* (pp. 3-23). Hoboken, NJ: John Wiley.

Declercq, K., & Verboven, F. (2010). Slaagkansen aan Vlaamse universiteiten: tijd om het beleid bij te sturen? *Impuls, 42*(2), 88-98.

De Fruyt, F., & Rolland, J. P. (2010). *PfPI. Beschrijving persoonlijkheid op het werk. Handleiding.* Amsterdam/Antwerpen: Pearson.

Duckworth, A., & Quinn, P. (2009). Development and validation of the Short Grit Scale (GRIT–S). *Journal of personality assessment, 91*(2), 166-174.

Duijnhouwer, H., Prins, F. J., & Stoking, K. M. (2012). Feedback providing improvement strategies and reflection on feedback use: Effects on students' writing motivation, process, and performance. *Learning and Instruction, 22*(3), 171-184. doi:10.1016/j.learninstruc.2011.10.003

EUROSTAT. (2017). Table: Annual expenditure on public and private educational institutions per pupil/student. http://ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=tps00067

Feldman, K., Smart, J., & Ethington, C. (1999). Major Field and Person-Environment Fit: Using Holland's theory to study change and stability of college students. *The Journal of Higher Education, 70*(6), 642-669. doi:10.2307/2649169

Feldman, D. C., & Whitcomb, K. M. (2005). The effects of framing vocational choices on young adults' sets of career options. *Career Development International, 10*(1), 7-25. doi:10.1108/13620430510577600

Flemish Ministry of Education and Training. (2008). Education in Flanders. A broad view of the Flemish educational landscape.*2013*(4/04). Retrieved from http://www.eui.eu/Documents/MWP/AcademicCareers/Countries/Belgium/BelgiumFlemishHigherEducation.pdf

Gati, I., & Asher, I. (2001). Prescreening, In-Depth Exploration, and Choice: From Decision Theory to Career Counseling Practice. *The Career Development Quarterly, 50*(2), 140-157. doi:10.1002/j.2161-0045.2001.tb00979.x

Germeijs, V., & De Boeck, P. (2003). Career indecision: Three factors from decision theory. *Journal of Vocational Behavior, 62*(1), 11-25. doi:10.1016/S0001-8791(02)00055-6

Germeijs, V., & Verschueren, K. (2007). High school students' career decision-making process: Consequences for choice implementation in higher education. *Journal of Vocational Behavior, 70*(2), 223-241. doi:10.1016/j.jvb.2006.10.004

Gottfredson, L. S. (1981). Circumscription and Compromise - A Developmental Theory of Occupational Aspirations. *Journal of Counseling Psychology, 28*(6), 545-579. doi:10.1037//0022-0167.28.6.545

Gottfredson, L. S. (1996). Gottfredson's theory of circumscription and compromise. In D. Brown & L. Brooks (Eds.), *Career choice and development* (Third ed., pp. 179-232). San Francisco: Jossey-Bass.

Grotevant, H. D., & Durrett, M. E. (1980). Occupational knowledge and career development in adolescence. *Journal of Vocational Behavior, 17*(2), 171-182. doi:10.1016/0001-8791(80)90002-0

Hattie, J. (2009). Visibly learning from reports: the validity of score reports. *Online Educational Research Journal.* Retrieved from http://www.oerj.org/View?action=viewPDF&paper=6

Holland, J. L. (1985). *Making vocational choices: A theory of vocational personalities and work environments* (Second ed.). Englewood Cliffs, NJ: Prentice Hall.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Lutz, FL: Psychological Assessment Resources Inc.

Kelchtermans, S., & Verboven, F. (2008). Participation and study decisions in a public system of higher education. *Journal of Applied Econometrics, 25*(3), 335-391.

Krieshok, T. S., Black, M. D., & McKay, R. A. (2009). Career decision making: The limits of rationality and the abundance of non-conscious processes. *Journal of Vocational Behavior, 75*(3), 275-290. doi:10.1016/j.jvb.2009.04.006

Lacante, M., VanEsbroeck, R., & DeVos, A. (2008). *Met een dynamische keuzebegeleidig naar een effectieve keuzebekwaamheid: tweede fase.* Retrieved from In opdracht van de Vlaamse Minister voor Onderwijs en Vorming, in het kader van het 'OBPWO-programma'. Brussel/Leuven:

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behavior research methods, 44*, 325-343. doi:10.3758/s13428-011-0146-0

Lipnevich, A. A., & Roberts, R. D. (2012). Noncognitive skills in education: Emerging research and applications in a variety of international contexts. *Learning and Individual Differences, 22*(2), 173-177. doi:10.1016/j.lindif.2011.11.016

McGrath, C., Henham, M., Corbett, A., Durazzi, N., Frearson, M., Janta, B., . . . Schweppenstedde, D. (2014). *Higher Education Entrance Qualifications and Exams in Europe: A Comparison.* Retrieved from http://www.rand.org/pubs/research_reports/RR574

Messick, S. (1990). *Validity of test interpretation and use*. Retrieved from Princeton, N.J.: http://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.1990.tb01343.x/abstract;jsessionid=85FF44E5AF52D2A6195D02D313377E7E.f03t02

Ministerie van Onderwijs en Vorming. (2009). Studiesucces Generatiestudenten in 2007-2008. Retrieved from http://www.ond.vlaanderen.be/hogeronderwijs/werken/studentadmin/studentengegevens/Studiesucces%20generatiestudenten%20in%202007.pdf

Müller, W. (2014). Educational inequality and social justice: Challenges for career guidance. *International Journal for Educational and Vocational Guidance, 14*(1), 21-33. doi:10.1007/s10775-014-9264-4

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational Interests and Performance: A Quantitative Summary of Over 60 Years of Research. *Perspectives on Psychological Science, 7*(4), 384-403. doi:10.1177/1745691612449021

OECD. (2003). Career guidance: new ways forward. In OECD (Ed.), *Education policy analysis*.

OECD. (2011). *Education at glance 2011: OECD indicators*: OECD Publishing.

OECD. (2014). *Education at glance 2014: OECD Indicators.* OECD Publishing.

OECD. (2016). *Education at glance 2014: OECD Indicators.*: OECD Publishing.

Oppedisano, V. (2009). *Open University Admission Policies and Drop Out Rates in Europe*. Retrieved from

Owen, S., & Froman, R. (1988). *Development of a college academic self-efficacy scale.* Paper presented at the National Council on Measurement in Education, New Orleans, LA.

Pitz, G. F., & Harren, V. A. (1980). An analysis of career decision making from the point of view of information processing and decision theory. *Journal of Vocational Behavior, 16*(3), 320-346. doi:10.1016/0001-8791(80)90059-7

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

Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288. doi:10.1037/0033-2909.130.2.261

Rounds, J., Smith, T., Hubert, L., Lewis, P., & Rivkin, D. (1999). Development of occupational interest profiles for O* NET. *National Center for O* NET Development, Raleigh, NC*.

Savickas, M. L. (2006). *Career construction theory*. Paper presented at the AACC 15th Annual National Conference, Sydney, Australia.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460-475. doi:10.1006/ceps.1994.1033

SweSAT. (2011). Swedish Scholastic Aptitiude Test. Retrieved from http://www.edusci.umu.se/english/swesat/

Tangney, J., Baumeister, R., & Boone, A. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271-322. doi:10.1111/j.0022-3506.2004.00263.x

UNESCO. (1997). *Operational Manual for ISCED 1997 (International Standard Classification of Education) (first edition)*. Retrieved from Paris: http://www.uis.unesco.org/Library/Documents/isced97-en.pdf

Van Daal, T., Coertjens, L., Delvaux, E., Donche, V., & Van Petegem, P. (2013). *Klaar voor hoger onderwijs of arbeidsmarkt? Longitudinaal onderzoek bij laatstejaarsleerlingen secundair onderwijs*. Antwerpen - Apeldoorn: Garant.

Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: The quality of motivation matters. *Journal of Educational Psychology, 101*(3), 671-688. doi:10.1037/a0015083

Vossensteyn, H., Stensaker, B., Kottmann, A., Hovdhaugen, E., Jongbloed, B., Wollscheid, S., . . . Cremonini, L. (2015). *Dropout and Completion in Higher Education in Europe*. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/library/study/2015/dropout-completion-he_en.pdf

Wessel, J. L., Ryan, A. M., & Oswald, F. L. (2008). The relationship between objective and perceived fit with academic major, adaptability, and major-related outcomes. *Journal of Vocational Behavior, 72*(3), 363-376. doi:10.1016/j.jvb.2007.11.003

Chapter 2: Technical manual: Practical implementation and criterion validity

Abstract

We start this chapter with a general description of practical implementation of the instruments. Next, we provide validation evidence for SIMON. In the following chapters, we will describe the basic principles and procedure of the development of SIMON. Yet, as more data is available each year, these principles have been and continue to be applied in larger samples of students. To fully understand our efforts to substantiate SIMON as a scientifically valid instrument to aid the choice of a higher education study program, we provide the available criterion validity information of SIMON-I and SIMON-C in the current chapter. We also elaborate on the important issue of test bias. More specifically, we evaluate test fairness with regards to gender, Socio-Economic Status (SES) and respondents with a different language background.

Description and use of the instruments

The main goal of this dissertation was to construct and substantiate an instrument that is ready for use for prospective students on the verge of choosing a study program. Moreover, apart from the instrument for this initial target group, SIMON is now also implemented post-enrollment. Both practical applications are described in what follows.

SIMON Pre-Enrollment

SIMON is freely available for all users through the website www.vraaghetaansimon.be. Users first make a personal profile in order to preserve accessibility of their results. When accessing the platform, they are free to decide what component they start with. Prospects who are uncertain about what program to choose usually start with SIMON-I. Users who know what programs they are interested in but are in need of information on their chances of success, can start with SIMON-C. Thus, because of the separate components and because the personal results are saved, prospective students can use and re-use SIMON at any point during their study choice process.

When completing the interest inventory, respondents receive their personal interests profile which consists of a graphical representation of the scores on the RIASEC dimensions and on the Academic scale (see Figure 1) together with a description of each of these scales. They can now also explore what programs match their personal interests. Users are allowed to retake the inventory whenever they please. Yet, they can also manually adapt scores on each of the dimensions after which the list of matching study programs is adapted on-the-fly. This feature was installed with the aim of letting users maximally explore educational programs and their features.

SIMON-C consists of all the skills and ability tests that have previously been validated (see Figure 2). Tests that are green have already been completed. Non-cognitive tests can be retaken at any time but cognitive test can only be retaken after 60 days. This limitation avoids
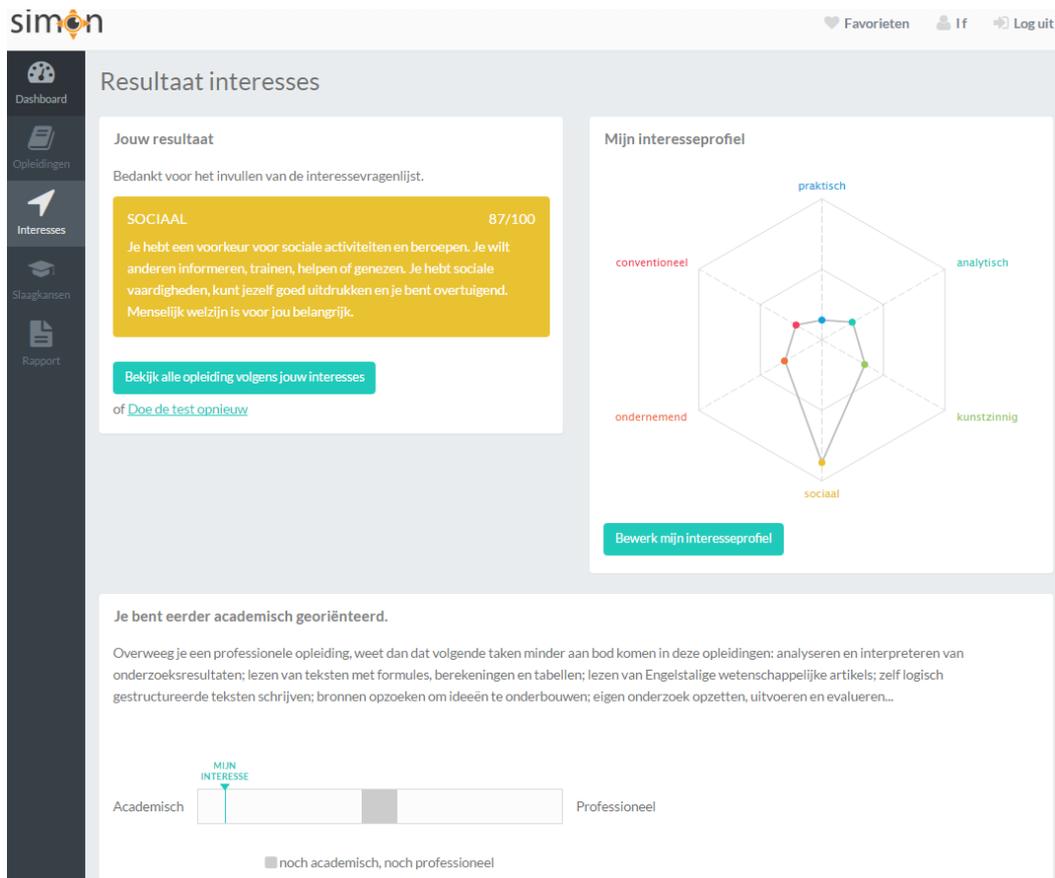
Figure 1. Screenshot of the output of SIMON-I. Users receive a graphical representation of their scores on the RIASEC dimensions and on the Academic scale.
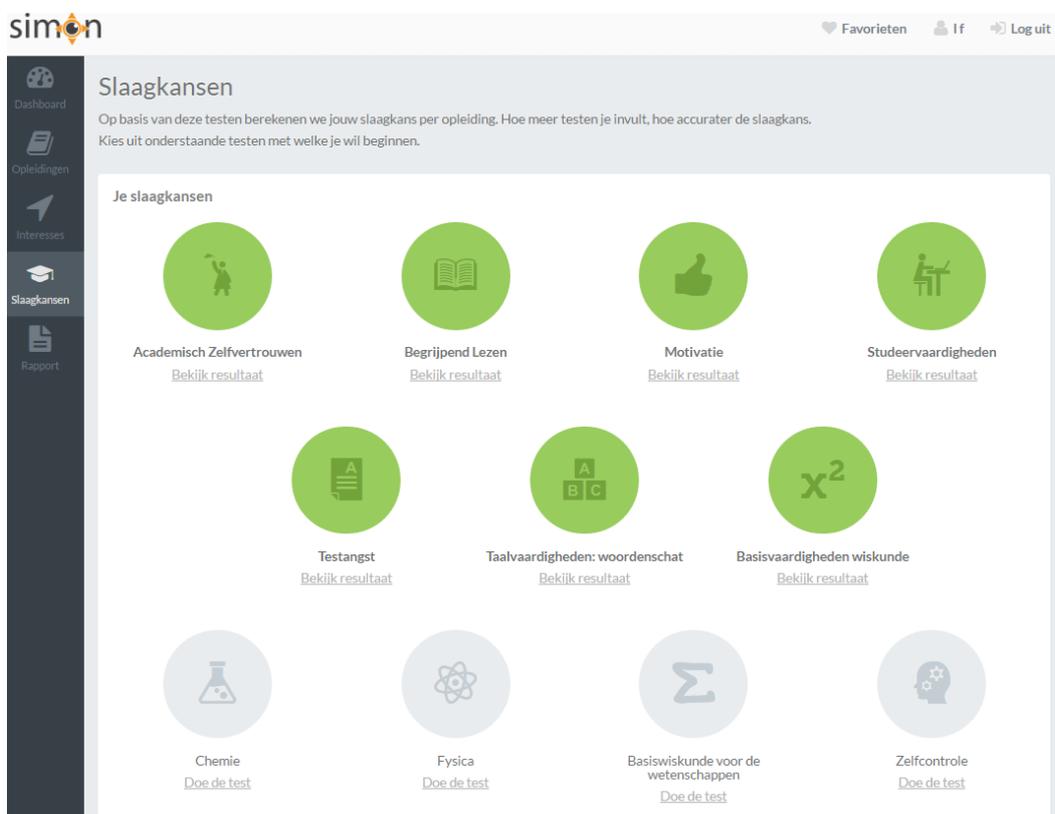
Fig. 2. Screenshot of the overview of SIMON-C tests.

users to keep retaking the test until they reach a maximal score, which would of course bias their prediction of success. When a test has been filled out, users first receive an explanation of why this skill is important in higher education. For example, the vocabulary knowledge test result starts with the statement "In higher education you are expected to convey your knowledge and ideas in an comprehensible way, for example when taking an exam, when writing a paper or when giving a presentation. This requires you to master a certain level of language skills. This test assesses your vocabulary knowledge". Users also get their personal score and information on the position of their score in relation to other prospective students. This information is shown both in numbers and through a graphical representation (see Figure 3).
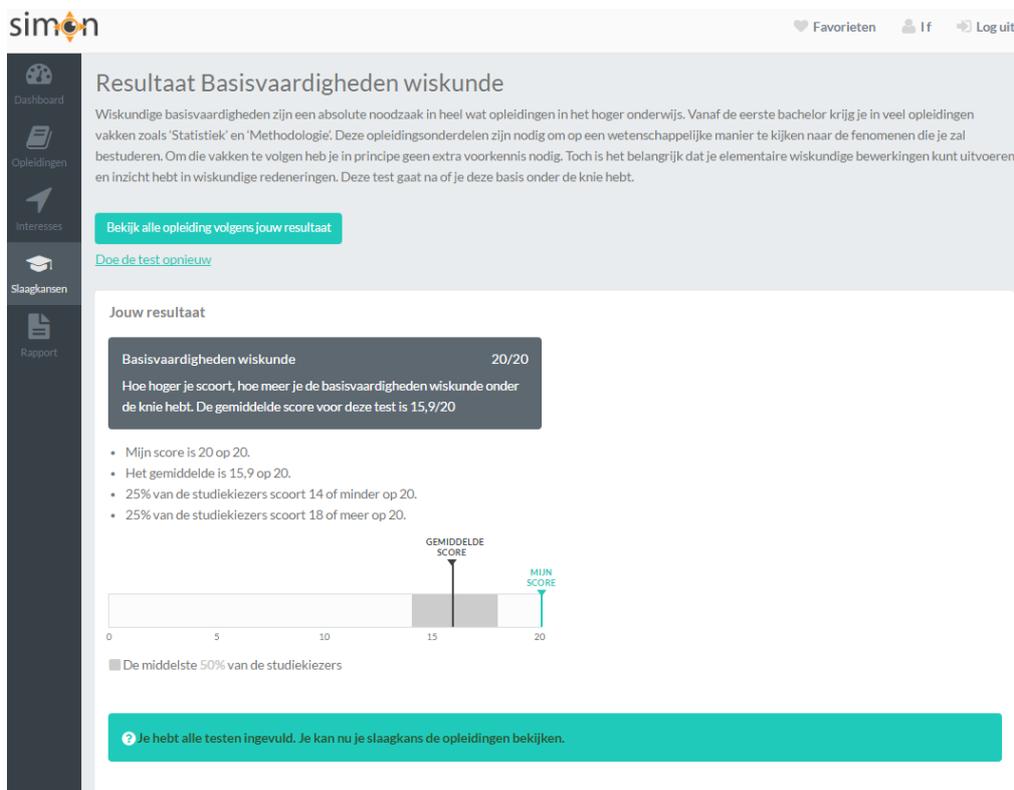
Fig. 3. Screenshot of the output of the SIMON-C basic mathematics test. Users receive an explanation of why this skill is important in higher education and they get their personal score and information on their score in relation to other users of the tool. This information is shown in numbers and through a graphical representation.

The results of SIMON-I and SIMON-C are integrated in the 'study program overview'. This page shows the user all included study programs and the match of these programs with the

personal interests and competencies. This match is expressed by a basic color code. Green indicates a good match (i.e., SIMON-I: the program matches one of the permutations of the three highest RIASEC scores; for SIMON-C: a high chance of passing as validated using the procedure described in chapter 1), red indicates there is no good match (i.e., SIMON-I: the program matches one of the permutations of the three lowest RIASEC scores; SIMON-C: low chance of passing) and orange means there is only a moderate match (i.e., all other programs). Because the list is quite extended (155 study programs are included), users can rank the programs according to their match with interests, their match with their competencies, the nature of the program (academic versus vocational) or all of these together.

To help the user in making a short list of programs, they can also check a 'favorites' box and decide to only view their favorites. An example of such a short list with information on the match with interests and competencies is shown in Figure 4.

When clicking on a specific program in the list, the users get program details page (see Figure 5). This page includes more information on the chances of success, and also a graphical representation of the interests of successful students in this program. This is complemented by the users' personal RIASEC graph which allows them to evaluate to what extent their interests do (or do not) correspond to the study program environment. Users also see what institutions organize the program and when clicking on the logo, they get redirected to the institutional program page which offers all information on the programs' contents and practicalities.

Users can also download or print their results (test scores and programs with information on their match to personal interests and competencies) in a report card.

Thus, SIMON is currently operational. Still, in constructing the instrument, possibilities for further expansion were incorporated. Components, tests, study programs and institutions can be added if wanted and validated. Each year the instrument is updated based on the data that was collected during the previous academic year.
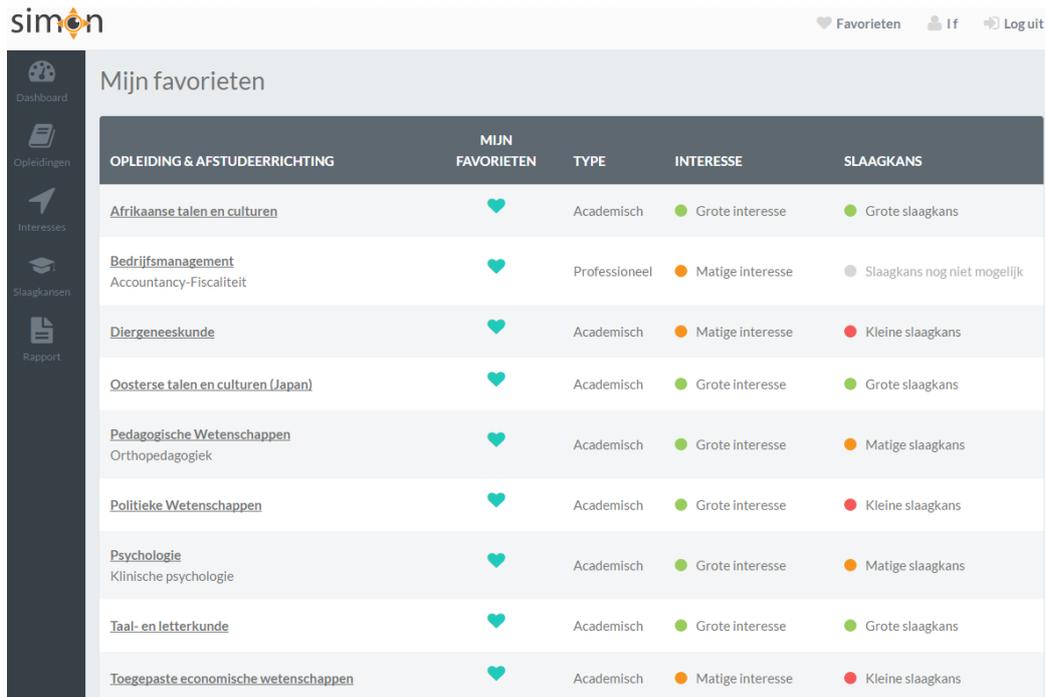
Fig. 4. Screenshot of SIMON (pre-enrollment) in which the respondent has indicated he/she only wanted to see his favorite study programs. The first column shows the name of the study program; the third showed whether this is a professional (vocational) or academic program; The fourth column shows the match between the respondents interests and the program (green: great interest, red: no interest, orange: moderate interest); and the last column shows the personal chance of success (green: high chance, red: low chance, orange: moderate chance/difficult to make predictions)
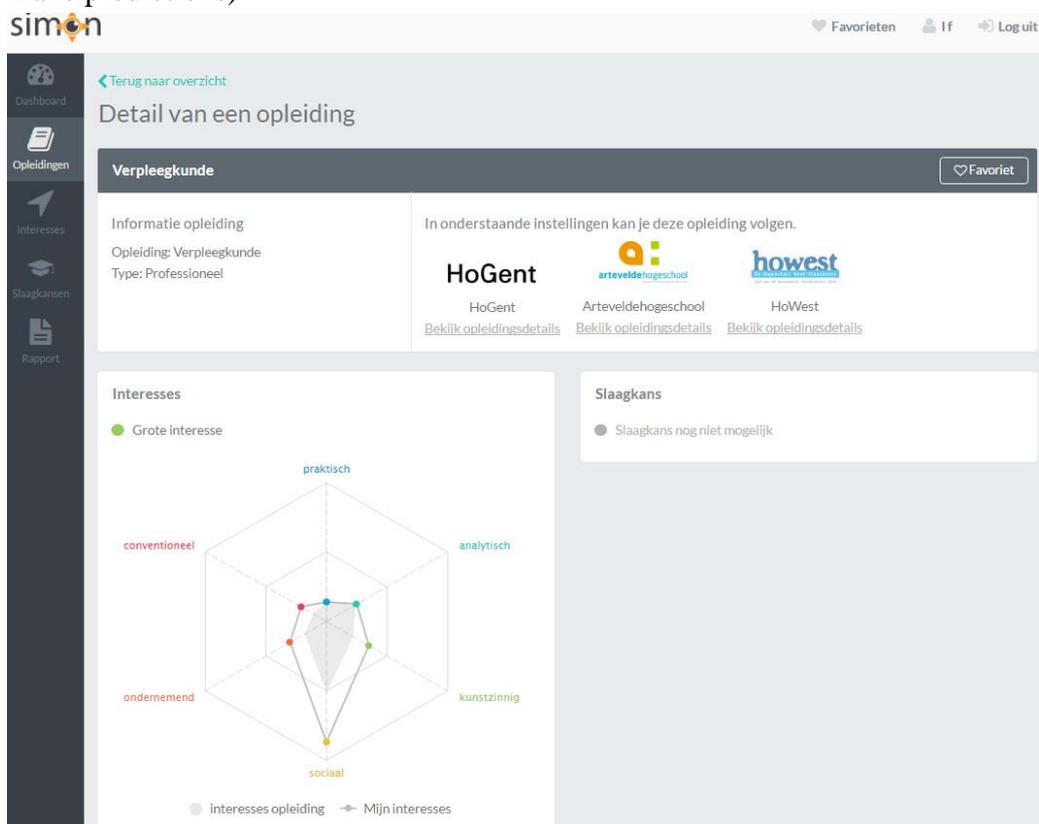
Fig. 5. Screenshot of a study program details page in SIMON

The website was launched in February of the year 2015 and during the first full school year ('15-'16) it was used by 20,000 unique prospects. In comparison, 49,487 students enrolled for the first time in higher education in Flanders in the academic year '15-'16 (Ministerie van Onderwijs en Vorming, 2016). These numbers demonstrate that prospective students were in need of a valid instrument that helps them in choosing a higher education study program.

SIMON as pioneer of the Columbus project

As a result of the success of SIMON, the Flemish government has decided to design a tool with the similar aim of aiding prospective students in their study program choice. In 2016, the Flemish so-called 'Columbus' project was launched. Whereas SIMON is limited to the co-operating institutions, the ambition of Columbus is to design an instrument that contains all study fields in Flanders. In doing so, several SIMON-modules (such as SIMON-I and the basic mathematics test) have been included in the pilot and validation phase of Columbus, together with language and non-cognitive tests that had been developed at other research institutions. The predictive validity of Columbus for academic and vocational programs across Flanders is currently being investigated. Thus, apart from being a ready-to-use instrument, SIMON has also played a pioneering role in the development of an instrument that may be applicable for prospective students across the Flemish region.

SIMON Post-Enrollment

A second practical value of SIMON is its' post-enrollment function. Newly enrolled students at Ghent University are invited to complete SIMON at the start of each academic year. Since the academic year '15-'16, participating students receive their results in a personalized feedback report (see the appendix in chapter 6 for an excerpt of an example report) three weeks into the academic year. This feedback report consists of three important pieces of information: 1. A personal chance of success (if validated following the principles described in chapter 1 and 5), 2. A personal score on each of the included tests together with information on the position

of the score in the cohort of students in the study program, and 3. For each of the assessed skills and abilities an overview of remedial activities that could improve the competency concerned. Interventions at both the central and the faculty level of the university are included.

Since the introduction of this application in 2015, 10,224 newly enrolled students at Ghent University have received a personalized feedback report, of whom 7,612 got a personal chance of success for the study program they enrolled in. Post-enrollment SIMON was specifically designed to identify students who lack the basic prerequisites and are thus at risk of failure. The feedback reports provide these students with a clear warning and give concrete and workable advice on how students can improve their skills and abilities. As we will demonstrate in chapter 6 and below, these reports do not fall short of their goal. They stimulate goal management strategies in students which can in turn improve their academic success and retention.

General features and criterion validity

Interests

The main goal of the SIMON interest inventory and its output was to maximize the exploration of (relevant) study options. In doing so, it was imperative that SIMON-I gives valid study advice and secondly, that it supplies the user with a manageable list of matching study options which broadens their view and encourages the in-depth exploration of viable options. The initial validation of SIMON-I will be described in chapter 3. Yet, our continuous data gathering allowed us to further evaluate whether SIMON-I meets this goal.

The assessment of whether SIMON-I can identify personally relevant study program options, can be accomplished using the concept of a 'hit rate'. This measure expresses what percentage of people in a given group would have been referred to that group by their interest scores. This approach involves classifying each successful student into one of Holland's six types on the basis of their study program. A student is counted as a hit if his or her highest

RIASEC score matches the first RIASEC letter of his or her study program. Thus, a student studying Social work (a predominantly S-program) would be counted as a hit if his or her highest average score was on the S-scale. If 50 out of 100 students would obtain high-point codes on scales that agree with their study program, the hit rate would be 50%. This approach provides quantitative evidence of validity based on the predominant interests of criterion groups (ACT, 2009). When unweighted hit rates are used with Holland-type criterion groups, the chance hit rate equals 17% (1/6) (ACT, 2009). At present, this SIMON-I first letter code agreement is 60.9% ($N = 5,883$). This is considerably higher than the 17% hit rate expected by chance and exceeds findings from widely used interest inventories such as the ACT interest inventory (i.e., values ranging between 31 and 55%, UNIACT, 2009). However, the SIMON-I matching algorithm and output is not limited to a first letter agreement. Following guidelines by Rosen, Holmberg, and Holland (1989), SIMON-I generates a list of matching programs based on all of the permutations of the personal RIASEC code. For example, when a respondent has S,E, and C as highest scores, all SE, SC, ES, EC and CE study programs are offered as a good match. This provides the greatest opportunity for successful exploration of study programs as no individual resembles only a single type. Using this algorithm, on average 86.2% of the successful students across 4 participating institutions ($N = 4,227$) would receive the study program that they are enrolled in as suggestion based on SIMON-I. Finding a 100% correspondence would be highly unlikely as interests themselves are not 100% stable. In their meta-analysis on the stability of interests, Low, Yoon, Roberts, and Rounds (2005) found that the estimated population correlation of interests at college age (18-21.9 years) was .67. Stability was also lower ($\rho = .58$) at the age of 16-17.9 years, which is the typical age Flemish students are required to choose their study program. As a result of these fluctuations, it is not unusual that a fraction of students graduate from a study program that does not or does no longer correspond to their personal interests. Also, other studies showed lower percentages of

correspondence between students' interest profile and their major. Harrington (2006), for example, found that 76% of the students graduated a major congruent with their Career decision-making system (CDM, Harrington & O'Shea, 1993) scores. Considering these findings, a correspondence of 86.2% can be considered good.

Another way of evaluating the congruence between RIASEC codes (for example between individual and study program codes) is by using the C-index (Brown & Gore, 1994). This index will also be described in chapter 3 and compares RIASEC codes based on the hexagonal distance between the letters. C ranges between 0 and 18, with higher scores indicating higher congruence. C is symmetrically and normally distributed, with a theoretical population mean of 9. Because our study programs are assigned two letter codes, we used the modified C-index as proposed by Eggerth and Andrew (2006). This modified C-index allows comparison between Holland code profiles of less than three letters in length and is obtained by sequentially comparing the first and second letters in both codes. Results showed that the agreement between individual codes and study program codes was significantly higher than the mean of 9 ($C = 14.34$, $t = 110.92$, $p < .001$). In comparison, Wessel et al. (2008) found a mean correspondence of 10.48 ($SD = 3.63$) between students' interests and college major using the Strong Interest Inventory.

Thus, the results of these correspondence analyses support that SIMON-I allows for identification of personally relevant study program options.

A second question concerns the ability of SIMON-I to broaden the view of users and to stimulate them to further explore viable study options. A sample of secondary education students ($N = 315$) was invited to evaluate this issue. 55.8% of the respondents said SIMON-I helped them in their choice process and 55.4% indicated that SIMON-I encouraged them to look into study options they had never even considered before. These numbers demonstrate that SIMON-I does aid study program choice and the (in-depth) exploration of options.

Competencies

The aim of SIMON-C was to provide users with a realistic appraisal of their chances of success in a specific study program. As explained in chapter 1, SIMON-C specifically intends to identify students who almost certainly lack the necessary skills to pass their first year of higher education. Thus, the focus was on a high prediction accuracy in the high-risk group, limiting false negative advice: only a small group of students gets a clear warning that a program is unattainable, but this prediction should be very accurate. As success predictions in SIMON are program-specific it is extremely difficult to give a proper estimate of the number of secondary education users that gets a negative advice (i.e., a low chance of passing). Yet, the available data of newly enrolling students at Ghent University ($N = 8,653$) showed that 10% received a low chance of passing for the program they enrolled in. Eventually, only 6% of them passed their first year. These students obtained on average 41% of their ECTS credits. Historical data showed that this corresponds to a chance of attaining the degree in 4 years (timely graduation of 3 years + 1 extra year) of 1.5%. In comparison, 70% of the students with a high SIMON-C chance actually passed. These students obtained on average 87% of their ECTS credits, which corresponds to a chance of attaining the degree in 4 years of 85% (see Table 1).

Table 1: Ghent University students academic years '12 - '13 through '15 – '16 ($N = 8,653$) study results by SIMON-C predicted chance of passing

SIMON-C chance of passing	Accuracy (% passing)	Average % ECTS credits obtained	Average chance of obtaining the degree (in 4 years)
Low	4%	41%	1.5%
Average	40%	68%	36%
High	70%	87%	85%

Post-enrollment SIMON: Feedback reports

To evaluate the effects of the post-enrollment feedback reports, 6,649 newly enrolled students were invited to complete a questionnaire in November 2016. 2,330 students started the

questionnaire and 1,849 completed it (27.8% of the newly enrolled students, 37.4% of the students who had received a report).

Results showed that a very high percentage of students (98.7%, $N = 1,983$) actually read the feedback report. The report was also well accepted: 77.5% found the feedback justified, 62.6% found it useful and 68% said they would recommend other students to complete SIMON. 54.2% of the total sample indicated that they were activated by the feedback report. Students with a low chance of passing were activated the most (see Figure 6): 31% of them actively participated in remedial activities, 28% said the report stimulated them to put more effort into their studies and 10% had considered changing study programs.
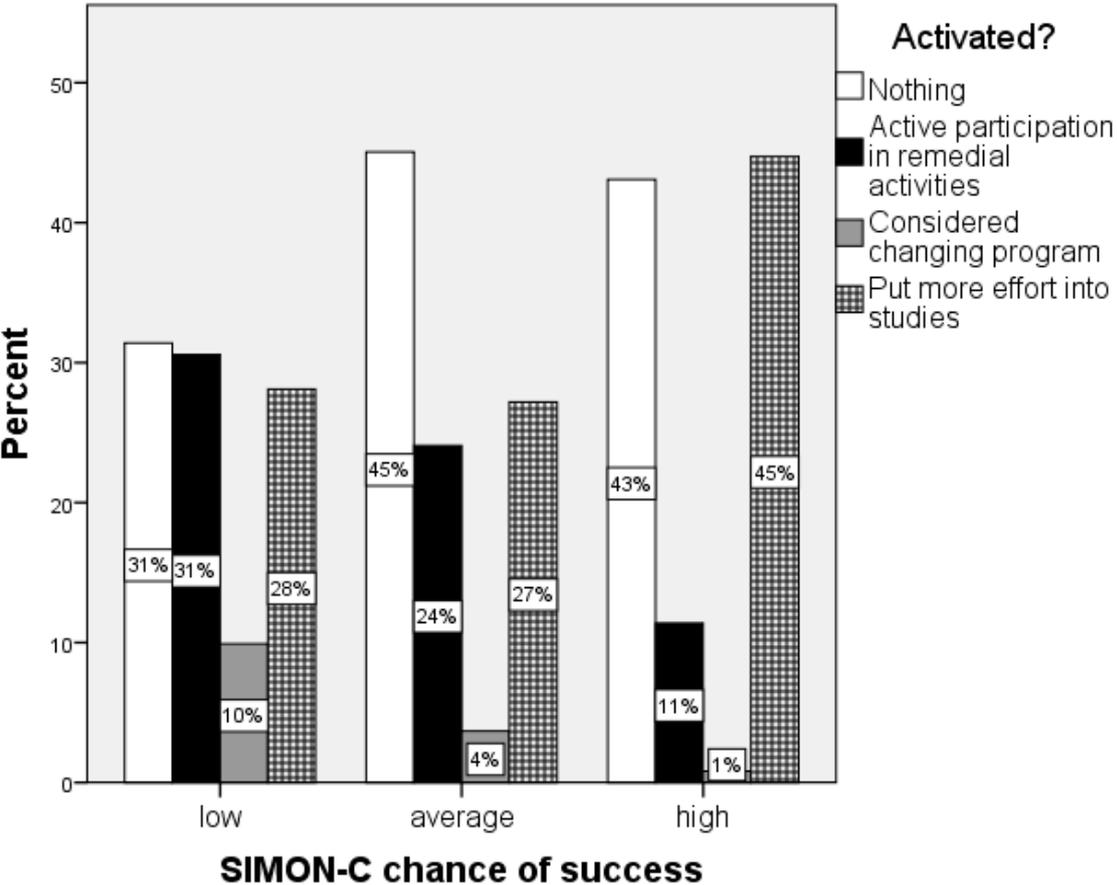
Fig. 6 Effect of post-enrollment feedback reports by SIMON predicted chance of success

When comparing the cohort of the academic year '14-'15 (when no feedback report was sent out to students) with the cohort of '15'-16 (when the first cohort of students received

feedback reports), data showed that students who had a low chance of passing and who had received feedback significantly more often (15%) changed study programs than students with a low chance of passing who did not receive feedback (9%) ($t = -2.60$, $p = .01$).

In sum, these results show that the feedback reports do generate the intended effect. Students who receive a feedback report reflect upon their starting competencies and on their match with the program they enrolled in. Especially students with a low chance of passing are encouraged to think about other (more fitting) study programs or to take up remedial courses in order to improve their chances of success.

Test bias and fairness

During the development of SIMON, we remained vigilant about bias against minority groups. A biased test is one that systematically over- or underestimates the value of the variable it is intended to assess, for a specific group. If this bias occurs as a function of a cultural variable, such as ethnicity or gender, cultural test bias is said to be present (Reynolds & Ramsay, 2003). To avoid this type of bias in SIMON, we evaluated fairness with respect to three important demographic variables: socio-economic status (SES), gender and different language background.

SES

First, it was important to evaluate the effect of SES on the SIMON-C predicted chance of success (Sackett et al., 2012). Paralleling the procedure of the Flemish Department of Education (Ministerie van Onderwijs en Vorming, 2012) to grant extra resources to schools with a high level of children low in SES, we used two indicators of SES: receiving a bursary and having a mother who did not attain a secondary education qualification. Students who met any of these criteria were categorized in the low SES group.

An important preliminary remark is that there were in fact significant differences in academic performance between low and high SES groups. Students with low SES attained a

lower percentage of their ECTS credits ($M = 57.5$ $SD = 37.6$) than other students ($M = 66.3$, $SD = 36.6$) ($t(15854) = -12.6$, $p < .001$). Students low in SES also passed significantly less often (26.8%) than other students (38.6%) (χ^2 (1, $N = 16,844$) = 180.9, $p < .001$). As a result, it would be plausible and perhaps even desirable that SIMON more often predicts a lower chance of passing to students low in SES. However, it would not be fair to (dis)advantage any group by disproportionally assigning low or high chances of passing. This fairness could be assessed by evaluating within groups of students who pass on the one hand and students who do not pass on the other whether students low in SES are more likely to receive a low or high chance of passing.

When looking at the group of students who passed, there was no significant relation between the SIMON prediction of success and SES group (χ^2 (2, $N = 3,292$) = .50, $p = .78$). There was however a significant relation within the group of students who did not pass (χ^2 (2, $N = 5,361$) = 8.3, $p = .02$). Thus, students with low SES more often received a low chance of passing, but this was justified on the basis of their lower performance. Students low in SES did not *unjustly* receive more negative advice than do other students. On the other hand, students low in SES were more often (17.4%) correctly identified as being at risk of failure than students who are not low in SES (14.3%). This means that students low in SES more often *correctly* received a warning that the study program of their choice was difficult to attain. In other words, SIMON did not (dis)advantage students with low SES by *wrongly* giving low or high chances of passing. Yet, SIMON did benefit low SES students by *correctly* signaling them that they were at risk of failure. Thus, these low SES risk students are encouraged more often to reconsider their program choice or to participate in remedial activities than other risk students. This may be a leverage for social equality in higher education, an issue which is taken up further in chapter 7.

Gender

The same procedure was applied to gender. Men passed less frequently (32.5%) than women (38.4%) (χ^2 (1, N = 16,698) = 61.34, $p < .001$). Men also obtained a significantly lower percentage of ECTS credits ($M = 60.3$, $SD = 37.9$) than did women ($M = 67.4$, $SD = 36.1$) (t(15708) = -11.7, $p < .001$). Consequently, SIMON more often gave a low chance of passing to men (10.9%) than to women (9.1%).

Yet, when singling out the students who passed, there was no significant relation between the SIMON chance of passing and gender (χ^2 (2, $N = 3,282$) = 1.1, $p = .57$). Neither was there within the group of students who did not pass (χ^2 (2, $N = 5,341$) = 1.4, $p = .49$). In other words: men more often received a low chance of passing, but this was justified on the basis of their lower performance. Men did not *wrongly* receive more negative advice than do woman. On the other hand, men were not more often *justly* identified as being at risk of failure than women. Thus, the assessment was functioning similarly for men and women.

With regards to SIMON-I, it was important to address gender differences in interests. Men and women are consistently found to differ in vocational interests, with men scoring higher on Realistic and Investigative interests and women favoring Artistic, Social and Conventional activities and occupations (see e.g., Su, Rounds, & Armstrong, 2009). The debate as to why men and women differ in their vocational interests is open, but some have suggested that these differences are an artifact of test construction. Therefore, it was important to take into account possible gender bias in the test construction phase. An important problem to address with regards to gender fairness is the differential functioning of items. Differential Item Functioning (DIF) occurs when respondents from different groups (in this case men and women) show differing probabilities of endorsing items after matching on the underlying trait that the item is intended to measure (Zumbo, 1999, p.12). For example, DIF would take place when women who have the same underlying level of R-interests as their male counterparts, would have a

lower probability of endorsing a specific R-item (a specific activity or occupation) less frequently than these men. In this case, the item would show bias towards men as men would be more likely to indicate they are interested in this specific activity or occupation. Wetzel, Hell, and Pässler (2012) showed that item response theory (IRT)-based DIF analyses were useful to address this issue.

We applied such a procedure, SIBTEST (Shealy & Stout, 1993), to the SIMON-I items. Results of the tests will be described in chapter 3 and show that 51% of the items showed bias. Importantly, with 47.6% of the bias items favoring women and 52.4% favoring men, there was no systematic bias against men or woman in any of the scales.

Different language background

With regards to language background, four different groups were considered: students with Dutch as native language (which is the language spoken in Flanders), students speaking French (which is the language spoken in the southern region of Belgium), students that speak a different EU-language, and students speaking a different non-EU language at home.

Students speaking Dutch passed significantly more often (40.5%) than students speaking French (27.4%), another EU language (20.6%) and another non-EU language (13.6%) (χ^2 (3, N = 10,974) = 131.21, p <.001). There were also significant differences in percentage ECTS credits obtained between language groups (F(3, 10507) = 71.86, p < .001)). These differences are shown in Table 2.

Within students who passed, there was no significant relation between the SIMON predicted chance of success and language background (χ^2 (6, N = 3142) = 3.34, p =.77). This means that the amount of students that *wrongly* received a low chance of passing is similar across language groups.

Table 2: Differences in obtained ECTS credits between language background groups

	N	Mean	*SD*
Dutch	9848	68,5	35,6
French	236	55,9	38,7
Other language (EU)	121	49,1	39.0
Other language (non-EU)	306	42,4	36,7
Total	10511	67,3	36,1

Within the group of students who did not pass there was a significant relation between the chance of success and language background (χ^2 (6, N = 5,110) = 17.92, p =.01). Students speaking Dutch more often received an average chance of success (83%) than students in any of the other language groups (78%, 75.3% and 75.7% respectively). Students with a different EU language more often received a high chance of passing (4.1% versus 2.9% for Dutch speaking, 1.5% for French speaking and 2.2% for other non-EU language speaking students). Thus, although the differences are small, students speaking a different EU language were more often *wrongly* classified as having a high chance of passing. This may indicate that these students master the basic skills to pass, but that they more often fail because of factors that are not assessed in SIMON.

In sum, the current evidence showed that SIMON does not (dis)advantage any groups and can thus be considered as fair.

Conclusion

In this chapter, we described how our research results are practically implemented, thereby demonstrating how the instrument can be used to aid (prospective) students in their choice of a higher education study program. We also presented more validity evidence. Results showed that SIMON can identify both personally relevant study options (SIMON-I) and students at risk of failure in their first year of higher education (SIMON-C). Moreover, evidence shows that SIMON did not exhibit bias towards specific (underrepresented) groups.

References

ACT. (2009). *ACT Interest inventory. Technical manual.* Retrieved from Iowa City: http://www.act.org/content/dam/act/unsecured/documents/ACTInterestInventoryTechnicalManual.pdf

Brown, S. D., & Gore, P. A. (1994). An evaluation of interest congruence indices: Distribution characteristics and measurement properties. *Journal of Vocational Behavior, 45*(3), 310-327. doi:10.1006/jvbe.1994.1038

Eggerth, D. E., & Andrew, M. E. (2006). Modifying the C index for use with Holland codes of unequal length. *Journal of Career Assessment, 14*(2), 267-275.

Harrington, T. F. (2006). A 20-year follow-up of the Harrington-O'Shea Career Decision-Making System. *Measurement and Evaluation in Counseling and Development, 38*(4), 198.

Harrington, T. F., & O'Shea, A. J. (1993). *The Harrington-O'Shea career decision-making system revised. Manual.* Retrieved from Circle Pines, Minnesota:

Low, K. S., Yoon, M., Roberts, B. W., & Rounds, J. (2005). The stability of vocational interests from early adolescence to middle adulthood: a quantitative review of longitudinal studies. *Psychological Bulletin, 131*(5), 713. doi:10.1037/0033-2909.131.5.713

Ministerie van Onderwijs en Vorming. (2012). *Omzendbrief SO/2012/01.* (SO/2012/01). Retrieved from http://data-onderwijs.vlaanderen.be/edulex/document.aspx?docid=14370.

Ministerie van Onderwijs en Vorming. (2016). Hoger Onderwijs in cijfers academiejaar 2015-2016. Retrieved from http://www.ond.vlaanderen.be/hogeronderwijs/werken/studentadmin/studentengegevens/HOIC_2015_2016.pdf

Reynolds, C. R., & Ramsay, M. C. (2003). Bias in psychological assessment: An empirical review and recommendations. In J. R. Graham & J. A. Naglieri (Eds.), *Handbook of psychology (vol 10): Assessment psychology* (pp. 67-93). Hoboken, NJ: Wiley & Sons.

Rosen, D., Holmberg, K., & Holland, J. L. (1989). *The College Majors Finder.* Odessa, FL: Psychological Assessment Resources, Inc.

Sackett, P. R., Kuncel, N. R., Beatty, A. S., Rigdon, J. L., Shen, W., & Kiger, T. B. (2012). The Role of Socioeconomic Status in SAT-Grade Relationships and in College Admissions Decisions. *Psychological science, 23*(9), 1000-1007. doi:10.1177/0956797612438732

Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika, 58*(2), 159-194. doi:10.1007/BF02294572

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychological Bulletin, 135*(6), 859-884. doi:10.1037/a0017364.

Wessel, J. L., Ryan, A. M., & Oswald, F. L. (2008). The relationship between objective and perceived fit with academic major, adaptability, and major-related outcomes. *Journal of Vocational Behavior, 72*(3), 363-376. doi:10.1016/j.jvb.2007.11.003

Wetzel, E., Hell, B., & Passler, K. (2012). Comparison of Different Test Construction Strategies in the Development of a Gender Fair Interest Inventory Using Verbs. *Journal of Career Assessment, 20*(1), 88-104. doi:10.1177/1069072711417166

Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores.* Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Chapter 3: Exploring vocational and academic fields of study: Development and validation of the Flemish SIMON Interest Inventory (SIMON-I)[1]

Abstract

A new, Holland-based Interest Inventory is proposed, intended to facilitate the transition from secondary to tertiary education. Specific interest items were designed to grasp activities that are prevalent during tertiary studies, including an Academic-track-scale to assist in the choice between academic and vocational-oriented programs. Interest profile descriptions are complemented by a list of matching study programs. Data from 3,962 students were analyzed to evaluate the underlying circumplex structure, the criterion validity of the Academic-track-scale and the study program RIASEC codes. It is concluded that the assessment and feedback tools are promising instruments to facilitate the transition to tertiary education.

Introduction

The Organisation for Economic Co-operation and Development (OECD; 2013) reported that 32% of tertiary students fail to graduate. In Flanders, the Dutch-speaking part of Belgium and the regional context for the present study, about 40% of university students succeed in terminating all courses successfully during the first year of tertiary education. These trends are alarming, even more so since first year performance is one of the best predictors of academic retention (de Koning, Loyens, Rikers, Smeets, & van der Molen, 2012; Murtaugh, Burns, & Schuster, 1999).

One of the critical aspects in preventing drop-out and improving success rates is adequate support and information during the study choice process. Students who carefully explore their options are more likely to end up in a program that suits their interests and potential, which in turn will lead to higher retention rates. For example, Germeijs and Verschueren (2007) showed that in-depth exploration of the environment during the study choice process led to a higher commitment to the chosen study program, which eventually resulted in better academic adjustment.

The exploration of personal interests is an important aspect of this self-investigation phase in the study choice process. Nye et al. (2012) showed in their meta-analysis that interests, and especially the fit between individuals and their environment, were strongly related to performance and persistence in academic contexts. It is thus important for people in the process of choosing a study program to explore both their interests and their study options, to end up in a matching program where drop-out will be less likely. In order to accomplish this daunting task, valid and accessible methodologies that encourage this self-exploratory process are required.

The Need for a New Interest Inventory

An abundance of interest inventories have already been developed, such as the widely used Self-Directed Search (Holland, 1985b) or the Strong Campbell Interest Inventory (Campbell, 1987). However, there are several reasons that may compel researchers and practitioners to create new instruments, particularly in the context of educational orientation.

First, most of the established interest inventories draw heavily or even exclusively on occupational titles to assess interest profiles. Yet, when students are asked to choose a field of study at the age of 17 or 18 (i.e., the age at which most students enroll in higher education), their ability to self-report on vocational interests through preferences for specific occupational titles may be constrained by a still limited understanding of how the world of work is organized (Grotevant & Durrett, 1980) and what is required in terms of knowledge and skills to adequately perform in different occupations. Moreover, when making this educational decision, students are more concerned with their level of interest in the respective fields of study than with the future job opportunities that might result from their study choice (Malgwi, Howe, & Burnaby, 2005). This is especially true since not all students end up in a job that matches their field of study (see e.g., Wolniak & Pascarella, 2005). It is thus essential that the matching of study programs to personal interests does not solely rely on job titles but also includes items that are related to specific activities prevalent in the study curriculum and practical training of college programs.

Second, most inventories have been developed and validated in the U.S. Since previous research has shown that cross-cultural application of interest inventories is not always without problems (Einarsdóttir, Rounds, Ægisdóttir, & Gerstein, 2002), there is a need for measures that are tailored to the specific regional context. In the current study context (i.e., Flanders), there is a pertinent lack of validated measures that link students' interests to the available higher education programs. Moreover, no tools are available that may aid students in making the

decision between pursuing an educational career at the academic or rather at the vocational level (see below).

Third, educational systems are organized substantially different across cultural, national, and even regional boundaries, and interest inventories and their feedback tools should be maximally aligned with these requirements at the institutional level. When making educational choices in Flanders, students need to decide on which study program they want to pursue at the end of secondary education (age 17-18), both in terms of study content or study field (e.g., engineering, law, psychology, foreign languages…) and study level (either academic or vocational track). Previous research has demonstrated that study *fields* can adequately be described and structured using well-established vocational interest models, like John Holland's (1997) RIASEC model (De Fruyt & Mervielde, 1996). The choice between study *levels* pertains to the difference between the academic track (organized by universities) and the vocational track (organized by colleges). This also corresponds to the distinction between tertiary-type A (or academic) and tertiary-type B (or vocational) programs as specified in the International Standard Classification of Education (UNESCO, 1997). While the focus in the latter is more on concrete and specialized professional skills and direct entry into the labor market, academic programs are more theoretical and research-oriented leading to a master degree. Moreover, students with an academic background typically occupy supervisory positions and work on more abstract and complex matters, whereas people graduating from vocational programs are more likely to work under supervision on concrete and specific tasks. With very few exceptions, study fields can be studied either at the theoretical or at the more applied level. For example, the academic Psychology program extensively studies the fundamental principles underlying human psychology, hereby considering different theoretical perspectives, whereas the vocationally oriented 'Applied Psychology' program focusses on the practical application of psychological principles. Most tertiary education students (39% of the population) across

OECD countries graduate from a type A program nowadays. Still, a significant group of 11% of the population graduates at tertiary-type B level. This proportion can reach as much as 29.67% (New Zealand) (OECD, 2014). Thus, a common shortcoming in existing interest measures is that these have little to say about which track, academic (tertiary-type A) or vocational (tertiary-type B), aligns best with a person's interest profile.

Finally, most inventories fall under copyright restrictions of test publishers (Armstrong, Allison, & Rounds, 2008) and are not publically available, which is a severely constraining factor for secondary education pupils on the verge of selecting a study program. Optimally, secondary education students should have easy and free access to reliable and validated assessment and feedback tools, encouraging the exploration of their interests and corresponding study programs.

The current paper describes the development of a new interest inventory that circumvents these problems. Specifically, the goal of this project is to develop an interest assessment inventory and accompanying feedback tool that is part of the broader SIMON (Study skills and Interest MONitor) project, a Flemish institutional initiative aimed at assisting secondary education pupils in their selection of a higher education program that maximally suits their interests and abilities. In this prospect, the newly developed interest inventory offers several advantages over previously developed scales such as the Self-Directed Search (SDS). Although both instruments aim to promote the exploration of interests in and by respondents using Holland's model as a guiding theoretical framework (see below), there are a number of important differences that should render this new instrument more appropriate to assess students' interests in the specific context described above. First, the new measure could be tailored to a distinct target audience of students in their final year of secondary education, who are all on the verge of selecting a higher education study program. For this target population, the ultimate objective of the interest-assessment consists of improving the match between their

personal interests and the available study programs, rather than obtaining a match between interests and work environments in general. As a consequence, the operationalization of this instrument should be different from the operationalization adopted when constructing traditional interest inventories, such as the SDS. Specifically, items will be constructed and selected that are a reflection of specific study activities in different programs, on top of the commonly included occupational titles. A second innovation is that this new assessment tool will also encompass an Academic scale, to help students in their choice between academic versus vocational programs.

Theoretical Background of the Newly Developed Interest Inventory

Holland's (1997) RIASEC interest model served as the guiding taxonomic framework for our new assessment instrument. Although not entirely free of debate and criticism (e.g., Furnham, 2001; Tinsley, 2000), the RIASEC framework is currently the most widely used and researched model to structure interest inventories around the world (Brown, 2002; Nauta, 2010). Central in Holland's theory is the assumption that both people and environments can be described in terms of their similarity with six different personality and environment types, i.e., Realistic, Investigative, Artistic, Social, Enterprising, and Conventional (for an excellent description of these types, see for instance Nye et al., 2012). The idea is that the degree of congruence or fit between a person and his or her environment significantly relates to higher levels of achievement and/or satisfaction. Moreover, the six theoretical types can be organized in a hexagonal structure, reflecting the level of psychological similarity between types. That is, adjacent types (e.g., Realistic and Investigative) are most strongly related whereas opposite types (e.g., Realistic and Social) are expected to show the least similarity. Prediger (1982) extended Holland's theory by showing that two dimensions underlie the interest circumplex, namely the People/Things and the Data/Ideas dimensions. In the People/Things dimension, the Things axis is anchored by the Realistic type while the opposite end of the dimension (People)

is anchored by Social. The Data-Ideas dimension has the Data axis intersecting the midpoint between Enterprising and Conventional and the Ideas axis intersecting the midpoint between Investigative and Artistic types (Rounds & Tracey, 1993).

Previous research has demonstrated that differences in vocational interests between university programs in Flanders are in accordance with Holland's theory (De Fruyt & Mervielde, 1996). Specifically, students in Industrial, Bio-agricultural and Applied engineering had the highest score on the Realistic scale. Students enrolled in Science and Bioengineering programs scored highest on the Investigative type. Language and History students had highest scores on the Artistic scale while students in Psychology and Educational sciences programs matched the Social type. Finally, Economy, Political/social sciences and Law students scored considerably higher on the Enterprising scale. Given the widespread acceptance of the Holland model, and its demonstrated relevance in the current context, i.e., the Flemish higher education system (De Fruyt & Mervielde, 1996), the RIASEC model seems particularly appropriate to serve as the conceptual basis of our new interest inventory. There are currently no inventories available in the Flemish community that are specifically designed to explore study interests according to the well-established Holland-model.

Academic Versus Vocational Study Programs

As a second innovation, we want the newly developed inventory to shed light on the often difficult choice between academic versus more vocationally oriented programs, because there are no specific requirements to enroll in either programs in Flanders. For this purpose, an academic-track-scale was introduced to assess a distinct interest dimension, here referred to as the 'Academic' factor. The idea is that within the six RIASEC interest types, this academic factor should differentiate between students who are more academically versus more vocationally oriented. This implies that students in the academic track share a common interest regardless of their field of study (and corresponding RIASEC profile) as opposed to students in

the vocational track. Since the focus in academic programs is more on theoretical and less on concrete professional skills, we expect these students to be more interested in specific academic study activities such as reading scientific literature and designing and conducting research.

An important issue in this regard concerns the relationship between the Academic scale and the existing six RIASEC interest scales. For instance, considerable overlap might be expected with Holland's Investigative type, as this type has a preference for activities such as abstract thinking and analyzing (Holland, 1997). Nevertheless, it is important to note that all fields of study (and corresponding interest types), including primarily Investigative programs can be studied at either type A or type B level (see e.g., OECD, 2011, Table 4.4). Even for Science programs, which are primarily Investigative (De Fruyt & Mervielde, 1996), there is the opportunity to choose between academic versus vocational tracks, and the numbers show that both options attract a considerable population of students. As such, the new Academic factor should not so much be seen as an additional (i.e., seventh) interest type; but rather as an additional interest dimension that differentiates between students within each of the six RIASEC types (and accompanying fields of study). In this regard, this dimension shows some resemblance to Holland's (Holland, 1985a) conception of level of training, and to Tracey and Rounds' (1996) idea of a prestige dimension. Specifically, Tracey and Rounds (1996) explain that the typical People/Things and Data/Idea dimensions can be thought of as orthogonal dimensions structuring the field of RIASEC dimensions, while the prestige dimension cuts through this interest circumplex adding a third and independent dimension. Hence, just as there are RIASEC occupational interests that can be sorted from low to high prestige, one can distinguish between RIASEC study interests that are either academic or rather vocational. Moreover, just as the prestige dimension shows some overlap with one of the primary RIASEC interest scales (i.e., Enterprising), the Academic factor can be expected to correlate with Investigative study interests.

Matching Interests to Study Programs

Helping students to identify fitting study programs is a two-step process where they (a) gain self-insight into their own study interests and (b) are informed about the interest profiles of the available study programs. Therefore, the newly developed interest inventory presented here is accompanied by a separate feedback tool that links the generated interest profiles to a list of congruent study programs. Importantly, the classification of environments, occupational or educational, in terms of Holland's RIASEC model is a challenging undertaking that can be approached from different angles. Prior work on the classification of environments has mainly focused on describing occupations in terms of the RIASEC dimensions, relying on three different procedures: the incumbent method, the empirical method and the judgment method (see, Rounds et al., 1999).

In the educational domain, however, conspicuously little attention has been devoted so far to the classification of study environments according to the RIASEC model (Reardon & Bullock, 2004). The current study extends the available literature in this domain by directly comparing the convergence between three different classification methods that can be applied to higher education study programs, i.e. (a) expert ratings, (b) students' mean interest scores, and (c) RIASEC descriptions of equivalent occupations (see further).

In the following section, an overview is given of the construction process that lead to the SIMON Interest Inventory, followed by a summary of the research purposes of the current study.

Construction and Initial Analysis of the SIMON Interest Inventory (SIMON-I)

In a first stage, an iterative procedure was used to generate the interest items for the new inventory. Items were constructed by three independent experts. Two of these experts can build on extensive experience in vocational interest assessment research, while the third expert has widespread knowledge in educational guidance and student counselling in particular. Items

were written to reflect a wide set of activities that are characteristic of the full range of tertiary educational programs organized in Flanders. Based on both original (Holland, 1985a, 1997) and more contemporary (Wille, De Fruyt, Dingemanse, & Vergauwe, 2015) descriptions of the six Holland interest types, these activities were subsequently grouped in accordance with the RIASEC framework. Finally, this set of educationally relevant activities was also supplemented with a list of occupational titles that can be liked or disliked. The choice of these occupational titles was inspired by prior taxonomic work in the Netherlands and Belgium on the positioning of professions within the RIASEC structure (Hogerheijde, Van Amstel, De Fruyt, & Mervielde, 1995; De Fruyt & Mervielde, 1996).

The initial item pool consisted of 173 items describing RIASEC activities (88 items) and occupations (85 items). In addition to the six Holland scales, a seventh scale was constructed to assess interest for academic (versus vocationally orientated) programs. Item generation for the 'Academic' scale was based on the assumption that pupils who want to enroll in an academic track must be interested in specific academic study activities, irrespective of their field of interest. Examples of such activities are *reading scientific literature* and *autonomous implementation and evaluation of research activities*. This resulted in a 12-item 'Academic' scale that intends to grasp a 'generic interest for the academic track' as opposed to more vocationally oriented programs. The initial questionnaire hence comprised 100 items measuring preferences for study activities and 85 items indicating occupational preferences.

Upon completion of the assessment module, test-takers would be presented a personalized interest profile summarizing the percentage scores on the six RIASEC scales and the Academic scale, supplemented with a list of matching study programs that they could consider. For this purpose, all available study programs were assigned a two-letter RIASEC code generated by experts in vocational interest assessment and relying on prior empirical work describing the distribution of RIASEC interests across study programs in Flanders (De Fruyt &

Mervielde, 1996). We used two-letter codes for study programs instead of the three-letter codes proposed by Holland. The main reason is that RIASEC codes for tertiary education programs in Flanders are yet to be empirically substantiated, and the use of more detailed three-letter codes for matching purposes would still be too audacious.

There was a high level of agreement across the three experts for over 95% of the study programs. In less than 5% of the programs, only two out of three experts assigned the same two-letter code, and for these cases a final code was assigned after deliberation. To give one example, the study program "Economy" was assigned the two-letter code "EC", reflecting the primarily Enterprising and Conventional nature of this field of study. Upon completion of the interest inventory, respondents would receive a list of all study programs that matched their personal interest profiles based on the new inventory. This matching procedure used the first three letters of the personal interest profile, linking this to all study programs that either shared the first two letters (irrespective of their sequence), or that had the first and the third letters in common. For example, a respondent with an AIRCES interest profile would receive study programs coded by experts as AI, IA and AR.

This first version, SIMON-I, was administered online to a sample of 295 secondary education students (age 17-18). Students were recruited from four secondary schools that offer a broad range of secondary education programs. Respondents were asked to fill out SIMON-I in the classroom under the supervision of teachers and to give extensive feedback. This feedback consisted of an overall rating (5-point Likert scale) indicating to what extent they agreed with the generated profile (i.e., the interest profile and the proposed study programs). They were also invited to highlight items that were difficult to interpret and to provide further qualitative feedback concerning the assessment.

Based on these data and feedback, a second version of the instrument was developed. In total, 30 of the original items were deleted because (a) they were easily misunderstood or (b)

they showed insufficient coherence with other items in the scale as evidenced by an increase in Cronbach's alpha coefficient when the items were deleted. Seven items did not have the highest correlation with the intended scale and were moved to the corresponding scales. For example, the occupational title 'communication manager' was initially included in the Social scale, though was relocated to the Enterprising scale given its empirical association with this interest dimension. Nine additional items were generated to obtain a more complete coverage of the study program portfolio.

The resulting version of the inventory (see Appendix for the English translation of the Dutch SIMON-I) comprised 98 activity items (11 Academic scale items and 87 RIASEC scales items) and 66 occupations (RIASEC scales items). Instructions were clear and concise: respondents were asked to indicate in a yes-no-format whether they would enjoy the activities and professions or not. We opted for a forced-choice format (yes-no) instead of a Likert scale because yes-no interest items are easy to score, quicker to administer and they are equally reliable (Dolnicar & Grun, 2007; Dolnicar, Grün, & Leisch, 2011). Scale scores were converted to range between 0 and 100 and indicated the proportion of 'yes' answers out of the total number of valid answers to both activity and occupation items. This second version served as the basis for further psychometric and structural evaluation.

Study Objectives

Having discussed the rationale and procedure behind the development of SIMON-I, the purpose of the current study is to provide initial evidence for the validity and practical value of this interest assessment tool in secondary education students on the verge of selecting a higher education program. To meet this aim, three central questions will be addressed.

First, given that SIMON-I is based on Holland's model of interests, an important question in the validation process concerns the structural validity of the proposed measure. Therefore, the internal consistencies of the RIASEC study interest scales and the presumed

circular structure of the Holland types (Holland, 1997) will be investigated first. Special attention will also be given to possible gender differences in item functioning (i.e., differential item functioning) and in structural validity.

Second, the current study aims to provide initial evidence for the criterion validity of the Academic-track-scale. Specifically, it will be examined whether this scale can adequately discriminate between students in academic versus vocational programs across and within fields of study. Further, given the anticipated overlap with the Investigative interest scale in particular, special attention in this validation process will be devoted to the issue of incremental validity.

Third, the current study presents a validation of the RIASEC study program codes that are used in the SIMON-I feedback tool. Specifically, the expert-rated RIASEC codes for study programs will be compared against (a) the mean interest scores of students in these study programs, further referred to as 'empirical program codes' and (b) occupationally-derived RIASEC codes, referred to as 'O*NET program codes' (see further). The idea behind the empirical program codes is that the interest profile of a study program can be derived from the average interest profile of people enrolled in this particular program. This approach is consistent with Holland's basic idea that the people constitute the environment, and has been used in previous research that attempted to characterize college environments (e.g., Harms, Roberts, & Winter, 2006). In order to have an additional check of the validity of the expert-rated program codes, we also incorporated occupationally-derived RIASEC codes for the study programs which were extracted from O*NET (e.g., for the program 'Clinical Psychology' we used the O*NET RIASEC code for the occupation of 'Clinical psychologist'). O*NET is a U.S. database that contains information on hundreds of occupation-specific descriptors, including RIASEC codes. O*NET ratings were validated by Rounds et al. (1999) and have been used in previous studies on the structural validity of Holland's RIASEC model, also outside the U.S. (Wille, Tracey, Feys, & De Fruyt, 2014). Recall that the O*NET database contains *occupational*

RIASEC profiles, and that the aim of the SIMON-project is to construct a measure that aids in the process of *study* choice. Hence, this approach explores the possibility of using occupation-level interest data to approximate the interest profile of corresponding study programs. A correspondence analysis of the three sources of program RIASEC codes (i.e., expert, empirical, and occupational) will be conducted for six different study programs. High correspondence of empirically obtained interest scores with O*NET job codes and experts' program codes would provide extra validity for the SIMON-I feedback module.

Finally, we will also evaluate the usefulness and face validity of the SIMON-I output by analyzing respondents' level of agreement with their feedback reports. Remember that this feedback report consisted of both an interest profile and a list of study programs that fit with this profile based on matching RIASEC letter codes.

Method

Procedure

SIMON-I was administered in an online Dutch version that automatically generated a feedback report consisting of an interest profile (RIASEC scale scores) and a list of corresponding higher education programs. Students across faculties and institutions were invited to fill out the inventory. Respondents were then invited to leave comments and to indicate their agreement with the received report (both the interest profile and the corresponding programs) on a scale from 1 to 5 ("To which extent do you agree with the generated profile?").

Participants

To be able to validate the output generated by SIMON-I, data from students in their last year of tertiary education were analyzed, based on the assumptions that students (a) gradually *gravitate* towards college majors that fit better with their interest profiles, and that (b) over time, students are also socialized in such a manner that study environments gradually reinforce and reward certain interest profiles. As a result of these two processes, students in their graduation

year are likely to have an interest profile that matches their program (Smart, 1997). Including data from students in their first years of education might distort the results since drop-out as a consequence of mismatch between interests and program is still probable at this stage. Thus, students in their final year of study were recruited across four different educational institutions (one of which offers academic programs and three of which offer programs in the vocational track). In total, 4588 higher education students accessed the assessment platform. Cases with more than 5% of missing values (nine items or more) were deleted, resulting in a final dataset of 3962 respondents. Of these respondents, 92.6% were enrolled in the academic track, 7.4% were enrolled in the vocational track and 68.5% were woman. In general, 50.8% of the student population in Flanders is enrolled in academic programs and 54.8% are woman (Ministerie van Onderwijs en Vorming, 2012), which means that our sample is more academic and more female than the general population of students. Given the nature of this research population (all students enrolled in their final year of tertiary education) we can be quite confident that the research participants are a homogeneous group of students aged between 21 and 23 years old.

Results

Structural validity

Descriptive statistics and internal consistencies of SIMON-I.

Table 1 shows the number of items and the internal consistency of the subscales. Cronbach's alphas in the sample ranged from .83 (Academic scale) to .93 (Social interest scale), which indicated good internal consistency. The underlying People/Things (P/T) and Data/Ideas (D/I) dimension scores were calculated according to the formula provided by Prediger (1982)[2]. This validated formula allows the transformation of RIASEC scores into two dimensions underlying the hexagonal structure of interests by using the Cartesian coordinates. The

---

[2] The People/Things dimension: (2*R)+(1*I)+(-1*A)+(-2*S)+(-1*E)+(1*C)
The Data/Ideas dimension: (0*R)+(-1.7*I)+(-1.7*A)+(0*S)+(1.7*E)+(1.7*C)

correlations between SIMON-I subscales and the underlying dimensions are presented in Table 2.

Evaluation of circumplex structure.

To evaluate the circular structure of the proposed RIASEC scales, both confirmatory factor analysis (CFA) (Browne's Covariance structure modelling approach, Browne, 1992) and randomization test of hypothesized order relations (RTOR) (Hubert & Arabie, 1987; Tracey & Rounds, 1993) were applied. The use of these two approaches to test circular structure is in accordance with suggestions by Nagy, Trautwein, and Lüdtke (2010), who also gave an excellent overview of the similarities and differences between these procedures. The circular structure was evaluated for the entire dataset and for men and woman separately.

The CFA tests of model fit were conducted using the CircE-package in R (Grassi, Luccio, & Di Blas, 2010). This package allows the implementation of Browne's approach and also provides a graphical representation of the results. The results of these structural analyses are shown in Table 3. For men, all fit indices indicated good fit of the data with the proposed circular model. For woman, results were mixed. RMSEA (>.08) indicated an unacceptable fit, while the other absolute fit indices SRMR (<.08) and AGFI (>.90) signaled a good fit of data with the proposed circular model. The incremental fit index CFI also indicated unacceptable fit (<.95). In the total sample, only RMSEA indicated unacceptable fit, all other indices showed good fit of the data with the circular model (Hooper, Coughlan, & Mullen, 2008; Tabachnick & Fidell, 2007). Thus, overall, results of CFA showed that the circular structure holds especially for men and for the entire sample. Furthermore, the spatial representation confirmed the theoretically expected RIASEC ordering in all samples, including the female sample.

Table 1 Descriptive statistics, internal consistencies and number of items of the SIMON interest inventory subscales

	Activities				Occupations				Total			
	N items	*M*	*SD*	α	*N* items	*M*	*SD*	α	*N* items	*M*	*SD*	α
R	14	16.97	22.51	.87	9	20.12	24.12	.79	23	18.21	21.87	.91
I	15	38.52	21.98	.74	14	27.50	24.37	.83	29	33.27	21.64	.88
A	13	34.62	26.79	.83	13	29.83	27.69	.86	26	32.35	25.95	.91
S	18	45.71	28.85	.89	10	37.26	29.67	.83	28	42.80	28.11	.93
E	13	57.91	39.61	.88	11	29.63	27.15	.83	24	37.88	27.42	.92
C	14	30.19	27.06	.86	9	18.19	23.06	.79	23	25.53	23.72	.90
Ac	-	-	-	-	-	-	-	-	11	53.20	30.10	.83

Table 2 Scale and dimension intercorrelations

	R	I	A	S	E	C	Ac	D/I	P/T
R	1	,48**	,17**	-,19**	,22**	,31**	,25**	-.03	.62**
I		1	,23**	,08**	,01	,15**	,58**	-.42**	.36**
A			1	,39**	,22**	-,02	,18**	-.46**	-.42**
S				1	,12**	-,04*	-,02	-.17**	-.78**
E					1	,66**	,20**	.67**	-.15**
C						1	,25**	.70**	.25**
Ac							1	-.10**	.13**
Data/Ideas								1	.09**
People/Things									1

Note. ** Correlation is significant at the 0.01 level (2-tailed).

D/I = Data/ideas. A negative correlation with D/I indicates a positive relation with the Ideas dimension.

P/T = People/Things. A negative correlation with P/T indicates a positive relation with the People dimension.

The software package RANDALL (Tracey, 1997) was used to conduct RTOR analyses. Holland's theory postulates that correlations between adjacent scales (e.g., R and I) should be higher than correlations between alternate scales (e.g., R and A) and correlations between opposing scales (e.g., R and S) should be lowest. This results in a total of 72 order predictions, and RTOR evaluates the percentage of predictions that are met based on the available data (Tracey & Rounds, 1993). The result of this test is commonly expressed by a correspondence index (CI) which varies between -1 (no order predictions were confirmed) to +1 (all order predictions were confirmed). Rounds and Tracey (1996) provide benchmarks (CI=.70 for U.S. samples and measures and CI=.48 for international contexts) to compare the magnitude of model-data fit. The results of the current study (see Table 4) indicated good model fit for the total sample (CI = .83, p = .017), as well as for men (CI = .97, p = .017) and woman (CI = .81, p = .017) separately. All CI values exceeded the U.S. benchmarks, which further substantiates that the data in all samples fit the circular order.

Table 3 Overview circumplex goodness of fit indices

	RMSEA	90% CI RMSEA	SRMR	AGFI	CFI	*df*	*p*
Men	.05	.04-.07	.02	.98	.99	2	20
Woman	.10	.09-.12	.06	.93	.94	3	20
Total	.10	.09-.11	.06	.93	.95	4	20

Note. RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardised root mean square residual; AGFI = adjusted goodness-of-fit statistic; CFI = Bentler comparative fit index; *df* = degrees of freedom; P = parameters.

Table 4 Randomization test of hypothesized order relations

Group	Predictions		Correspondence Index	p
	Met	Tied		
Men	70	2	.97	.017
Woman	65	0	.81	.017
Total	66	0	.83	.017

Gender differences.

As previous research on Holland's interest dimensions has systematically shown gender differences in RIASEC interest scores (Su et al., 2009), specific attention was given to these differences and to the possible occurrence of gender bias in the developed scales. To establish whether there is an overall effect of gender on interest scores, discriminant analysis was used because of the interdependence of interest dimensions. In this analysis, all seven interest scales were considered simultaneously. The analysis was complemented with univariate tests to specify the contribution of each interest type, as advised by Borgen and Seling (1978). Discriminant analysis indicated that overall, there are gender differences in scale scores (Wilks' Lambda = .698, Chi square (7) = 1423.71, $p < .01$)). Independent samples t-tests showed gender differences on all seven scales (see Table 5). Specifically, men scored higher on Realistic, Investigative, Enterprising and Conventional interests while woman favored Artistic and Social interest dimensions. Men also obtained a higher score on the Academic scale compared to women. The two largest differences between men and women were found for Social and Realistic interests (Cohen's $d = -.93$ and Cohen's $d = .86$ respectively). This gave rise to a large effect size of 1.06 for the underlying P/T dimension. Men and women also differed on the D/I dimension, albeit to a lesser extent (Cohen's $d = .40$).

Table 5 Gender mean differences in interests: mean, standard deviation, univariate F and Cohen's d

	Men *M* (*SD*)	Woman *M* (*SD*)	*F* (1,3960)	*d*
R	31.03 (25.78)	12.30 (16.79)	489.78*	0.86
I	36.76 (23.15)	31.66 (20.72)	28.52*	0.23
A	27.55 (24.05)	34.56 (26.49)	33.21*	-0.28
S	26.90 (22.48)	50.14 (27.40)	148.27*	-0.93
E	45.21 (28.42)	34.50 (26.28)	21.61*	0.39
C	32.22 (25.51)	22.45 (22.18)	81.27*	0.41
Ac	61.27 (28.92)	49.47 (29.91)	7.35*	0.40
Data/Ideas	22.30 (101.53)	-15.78 (89.21)	59.32*	0.40
People/Things	4.49 (90.72)	-90.63 (88.57)	.89*	-1.06

* p < .001

Differential item functioning (DIF) was tested to investigate the extent to which the observed gender differences reflect a real difference between men and woman or whether they are an effect of gender bias in the items of the newly constructed scales. SIBTEST (Shealy & Stout, 1993) was used for this purpose, which is an item response theory based procedure. In this approach, a so called valid subtest is used as an estimate of the target trait being measured and the DIF test evaluates how the items differ in their performance in the two groups that are being compared by conditioning them on the trait level of the examinees. The procedure examines whether the resulting DIF statistic (β) is significantly different (*p* < .001) from 0 and which group (men or women) is being favored (Einarsdóttir & Rounds, 2009). The results in Table 6 indicate that half of the interest items showed significant DIF.

Importantly, for the Investigative, Artistic, Social, Enterprising, Conventional and Academic interest scales, there is an approximately equal number of items that favor men and women. For the Realistic scale, 4 items favor women as opposed to 8 items favoring men. Concerning the overall level of DIF, it can be noted that only the Investigative scale has beta values that are considered high (> .200 as in Einarsdóttir and Rounds (2009)).This indicates

that although there is gender bias in many of the interest items, this bias does not systematically affect the interest scale scores of one specific group.

Table 6 Number and percentage of items showing DIF for the SIMON-I scales

	Scale N	Differential item functioning (ß)					
		N items showing DIF	% items showing DIF	Favor women		Favor men	
				N	M (ß)	N	M (ß)
R	23	12	52	4	0,161	8	-0,127
I	29	14	48	7	0,255	7	-0.252
A	26	19	73	9	0,132	10	-0,121
S	28	11	39	5	0.115	6	-0.124
E	24	12	50	7	0.125	5	-0.138
C	23	9	39	4	0.128	5	-0,099
Ac	11	7	64	4	0.124	3	-0,160

Criterion validity of the Academic scale

Validation evidence for the 'Academic' scale was obtained by comparing mean scores on this scale between students enrolled in academic programs with those of students in vocational programs. We conducted these comparisons both across and within different fields of study.

To check whether the 'Academic' scale differentiates between academic or vocational interests *across* fields of study, an independent samples *t*-test was performed to assess a global difference in academic interests between respondents from academic programs and those enrolled in vocational programs. A significant difference was observed, $t(3960) = 8.40$, $p < .001$. Students in the academic track had a mean score of 54.24 ($SD = 30.05$), while students in the vocational track scored on average 40.06 ($SD = 27.55$), indicating that the Academic scale was able to differentiate between students in the academic and in the vocational track (Cohen's

d = .49). By way of comparison, Cohen's *d* effect sizes for the six RIASEC scales were -.16, .46, .05, .38, .00 and -.10 respectively.

Because of the relatively high correlation between the Academic scale and the Investigative scale (r = .58, p < .001), we performed additional analyses to substantiate the added value of this newly developed scale. To this aim, we first performed a new independent samples *t*-test on the subgroup of respondents whose primary (i.e., highest) interest score was Investigative. The results indicated that even within this subgroup, the Academic scale was able to differentiate between students in an academic (*M* = 68.21, *SD* = 26.82) versus a vocational (*M* = 54.00, *SD* = 24.56) track ($t(716)$ = 3.14, p < .01). Second, a hierarchical logistic regression analysis predicting the likelihood of enrollment in either the vocational or the academic track, showed incremental predictive validity of the Academic scale over and above Investigative scale scores ($\chi^2(1)$ = 16.65, p < .001).

Similarly, independent samples *t*-tests *within* the same field of study showed significant differences in Academic scale scores between students enrolled in academic versus those enrolled in vocational programs. For example, 'Chemistry' is offered both as a type A and a type B program. Although both programs share the same RIASEC program code (i.e., 'IR'), there is a significant difference in scores on the Academic scale between students enrolled in the academic track (*M* = 76.19, *SD* = 17.65) and those enrolled in the vocational track (*M* = 45.38, *SD* = 27.69) ($t(33)$ = -3.79, p < .01). To give another example, a similar difference was found between students from the academic 'Economical Sciences' program (*M* = 81.36, *SD* = 25.36) and those from the vocational 'Company Management' program (*M* = 45.12, *SD* = 26.38) , ($t(71)$ = -5.45, p < .001), despite their corresponding RIASEC interest code (i.e., 'EC').

Output evaluation

Correspondence analysis of study program RIASEC codes.

The expert-rated study program RIASEC codes that complement the SIMON-I interest profiles were validated by investigating their level of correspondence with (a) the mean RIASEC interest scores of respondents enrolled in these programs (i.e., the empirical program codes) and (b) the O*NET RIASEC codes of occupations corresponding to these study programs. Note that this analysis was restricted to a set of six different study programs that were selected based on (a) the theoretical positioning across the interest circumplex (i.e., one program for each of the six key points of the Holland interest hexagon) and (b) on the highest response rates within the respective interest types: Civil Engineering (Realistic), Bioengineering (Investigative), Languages (Artistic), Clinical Psychology (Social), Economy (Enterprising) and Medical and Health Care Management (Conventional). Corresponding job titles (i.e., biochemical engineer, civil engineer, clinical psychologist, interpreters and translators, economist and medical and health care manager) were searched through O*NET Online and the RIASEC codes for these job titles were retrieved from the O*NET database. To make the comparison between these three corresponding study program codes (i.e., expert-rated, empirical, occupational), a range of congruence indices are available (see e.g., Spokane, 1985). For the present study, we chose to use the C-index (Brown & Gore, 1994) because of (i) its consistency with Holland's theory, (ii) its normal distribution, and (iii) the ease of calculation and interpretation. Since experts assigned two-letter codes to programs as opposed to three-letter codes in the O*NET database, we use the modified C-index as proposed by Eggerth and Andrew (2006). This modified C-index allows comparison between Holland code profiles of less than three letters in length and is obtained by sequentially comparing the first and second letters in both codes. Comparison is based on the hexagonal distance between the letters. C ranges between 0 and 18, with higher scores indicating higher congruence. C is symmetrically

and normally distributed, with a theoretical population mean of 9. Table 7 summarizes the results of this correspondence analysis.

Before comparing the empirical program codes with O*NET and expert-generated codes, we inspected the mean SIMON-I interest scores across the six study programs. A one-way multivariate analysis of variance (MANOVA) indicated that, in general, students within a certain program indeed scored higher on the interest domain that corresponds with the theoretical position of that program in the hexagon. Specifically, post-hoc Tukey tests indicated that the highest score on Realistic was found for Civil Engineering ($F(5,1479) = 122.10$, p <.001); the highest score on Investigative for Bioengineering ($F(5,1479) = 83.14$, p <.001); the highest score on Artistic was for Languages ($F(5,1479) = 72.74$, p <.001); the highest score on Social was for Clinical Psychology ($F(5,1479) = 214.05$, p <.001) and the highest score on Enterprising was for Economy programs ($F(5,1479) = 146.76$, p <.001). There was only one exception: students in Health Care Management and Policy had higher Conventional scores than students from four programs, but lower scores than students from Economy programs ($F(5,1479) = 251.07$, p <.001). In general, these results indicate that SIMON-I meaningfully differentiates between students from theoretically different fields of study.

The agreement between O*NET codes and empirical program codes was significantly higher than the mean of 9 for Languages programs (C = 13.97, $t = 20.69$, $p < .001$), Health Care Management and Policy (C = 12.48, $t = 9.80$, $p < .001$), Bioengineering programs (C = 11.56, $t = 8.49$, $p < .001$) and Civil Engineering (C = 10.95, $t = 4.35$, $p < .001$). The agreement with O*NET RIASEC codes was not significantly different from the mean for Clinical Psychology (C = 8.93, $t = -.42$, $p = .67$). For Economy programs (C = 7, $t = 14.21$, $p < .001$), the agreement was lower than the mean. Since O*NET contains occupational information whereas expert codes were specifically given with study programs in mind, we expected the overall congruence with experts' ratings to be higher. This was confirmed: All C-indexes comparing empirical with

Table 7 Comparison of mean SIMON RIASEC codes across study programs with O*NET RIASEC codes

Theoretical position	O*net job title	N	Field of study/Major	R M (SD)	I M (SD)	A M (SD)	S M (SD)	E M (SD)	C M (SD)	O*NET Code	O*NET Mean C-index	Experts Code	Experts Mean C-index
R	Civil engineer	82	Civil Engineering	49.98 (24.75)	41.16 (23.66)	25.66 (24.25)	19.46 (20.93)	40.16 (24.96)	31.62 (24.15)	RI	10.95*	IR	10.54*
I	Biochemical engineer	224	Bioengineering	38.40 (24.4)	52.82 (19.24)	24.69 (22.42)	25.04 (22.19)	35.72 (24.72)	26.54 (20.76)	IR	11.56*	IR	11.56*
A	Interpreters and translators	233	Languages: Interpreter, translator, multi-Linguistic Communication and Languages	8.06 (12.51)	23.64 (17.96)	55.74 (22.83)	46.18 (24.34)	33.87 (24.66)	14.06 (15.77)	AS	13.97*	AS	13.97*
S	Clinical Psychologist	351	Clinical Psychology	8.71 (12.57)	30.82 (19.90)	40.55 (26.90)	69.48 (18.73)	25.78 (22.00)	11.08 (13.60)	IS	8.93	SA	14.83*
E	Economist	475	Economy: Applied Economic Sciences, Economic Sciences, Business Administration	22.64 (22.59)	23.88 (17.74)	23.69 (21.82)	25.93 (22.21)	65.06 (21.97)	53.66 (22.58)	IC	7.00	EC	14.07*
C	Medical and health care managers	120	Health Care Management and Policy	16.40 (19.70)	34.59 (21.05)	27.77 (24.78)	52.72 (23.55)	57.51 (22.21)	48.12 (24.07)	EC	12.48*	SE	12.14*

Note: * C-index significantly higher than the population mean (which is 9 as the C-index is normally distributed).

expert codes were significantly higher than the mean of 9. Bioengineering and Language programs were assigned the same letter code by experts as by O*NET, and had thus the same C-index. Civil Engineering programs showed slightly lower congruence with the experts code than with the O*NET code ($C = 10.54$, $t = 3.49$, $p < .001$). Health Care Management and Policy programs ($C = 12.14$, $t = 10.13$, $p < .001$), Economy programs ($C = 14.07$, $t = 29.42$, $p < .001$) and Clinical Psychology programs ($C = 14.83$, $t = 36.41$, $p < .001$) had higher congruence with experts' ratings.

A one-way multivariate analysis of variance (MANOVA) was conducted to test the hypothesis that respondents from different programs scored significantly higher on the interest scale that corresponds to the theoretical position of their study program on the circumplex. In other words: do Civil Engineering students score higher on the Realistic scale, Bioengineering students higher on the Investigative scale and so on. Results and post hoc-tests confirmed this, with the exception of Health Care Management and Policy in which respondents scored significantly higher on the Conventional scale than students from four other programs, but lower than respondents from Economy programs.

Respondent agreement with suggested feedback.

Respondents were asked to indicate their agreement with the received results (both the interests' profile and the corresponding programs) on a 1 to 5 scale (ranging from "*I completely disagree*" to "*I completely agree*"). 1367 respondents indicated their agreement with the interest profile and 1358 respondents evaluated the suggested study programs. 75% of the respondents indicated to agree with the presented interest profile (score 4 or 5); 16.2% agreed moderately (score 3) and only 8.8% did not agree (score 1 or 2) with this part of the feedback report. The mean agreement score was 3.86 ($SD = .91$). Regarding the proposed study programs, 55.5% of the respondents agreed (score 4 or 5); 21.9% agreed moderately (score 3); and 22.6% did not agree (score 1 or 2) with their feedback report. The mean agreement score was 3.41 (*SD*

= 1.08). We also explored whether these agreement scores were related to any of the interest scales measured by SIMON-I. Students who scored highest on the Social scale were most satisfied ($M = 4.07$); while those who scored highest on the Realistic scale were least satisfied ($M = 3.67$) with their interest profile. Agreement with the proposed study programs was not related to any of the interest scores. Overall, these results indicate that most respondents tended to agree with the profile that was generated by SIMON-I.

Discussion

The aim of this study was to document and validate SIMON-I and its feedback tool. SIMON-I is a new and freely available interest measure tailored to a target audience of students on the verge of selecting a higher education study program. The accompanying feedback tool aims to facilitate this choice process by providing respondents with a list of study programs that are matched to their interest profiles. SIMON-I also introduces a new academic-track-scale that deals with the often difficult choice between academic (tertiary-type A) versus vocational (tertiary-type B) programs. Overall, our findings speak for the validity and usefulness of SIMON-I and its feedback tool in the context of educational guidance and counseling.

One of the features that makes Holland's interest model so appealing for test developers pertains to its structural assumptions (Nauta, 2010). Specifically, the well-defined position of the six personality and environment types across the interest circle enables the analysis of person-environment congruence, an element that is highly relevant for both career researchers and practitioners. The structural validity of SIMON-I was confirmed in the present study by evaluating the underlying circumplex structure using both CFA and RTOR. With CFA, several fit indices showed a good fit of the data with the circular ordering, especially in men. RTOR revealed good fit of the data with the circular structure in all samples.

Our findings regarding gender differences in interest scores are largely in line with those reported by Su et al. (2009), showing that men generally scored higher on Realistic and

Investigative interests compared to women obtaining higher scores on Artistic and Social interests. Two findings in the present study, however, diverged from Su and colleagues' meta-analysis. First, SIMON-I did not reveal significant gender differences for Conventional interests, while Su et al. (2009) found women to score significantly higher on this scale. Second, contrary to the null findings reported by Su et al. (2009), SIMON-I did reveal significant differences between men and women in terms of Enterprising interests (i.e., men scoring higher), reaffirming earlier work in this area (e.g., Betz & Fitzgerald, 1987). Consistent with Su et al. (2009), the largest gender differences were found for the 'People/Things' dimension with men favoring working with things and women preferring working with people. These findings can further be considered in the context of a broader field of research dedicated to the structural (in)variance of interest models across gender. Previous studies on this topic have been inconclusive (Beinicke, Pässler, & Hell, 2014), with some reporting structural invariance across gender (e.g., Darcy & Tracey, 2007; Nagy et al., 2010), and others providing evidence for gender differences in the underlying structure of interests (e.g., Hansen, Collins, Swanson, & Fouad, 1993). In the present study, even in women, several fit indices showed good fit of the data with the circular model. Also, the spatial representations confirmed the theoretically expected RIASEC ordering. Moreover, the CI values found with RTOR in this study exceeded U.S. benchmarks and CI values established previously in a Flemish population of higher education students that were assessed with a translation of Holland's Self-Directed Search. Specifically, Wille et al. (2014) used the Dutch authorized adaptation of the SDS to measure vocational interests in final year higher education students and observed a CI of .69 for this instrument. This could suggest that SIMON-I, with a CI of .83 for the total sample, is better at capturing the circular order of interests compared to the SDS in Flemish higher education students.

Findings regarding gender differences in interest scales also raise the question of gender fairness in interest inventories (Pässler, Beinicke, & Hell, 2014). The results of differential item functioning tests performed on SIMON-I indicated that many of the interest items indeed showed bias. Nevertheless, this bias was not systematically directed against either men or women in any of the scales. We are aware of the potential problems associated with confirming gender differences as a result of gender-biased interest scales. Therefore, data from additional samples will be used in future work to further explore whether there is a need to replace items (especially in the Realistic and Investigative domains) to obtain more gender-fair interest scales.

The process of matching people to environments based on their interest profiles requires not only that personal interests are mapped (e.g., using an interest inventory) but also that environments are summarized in terms of their most prevalent interest-related characteristics. One of the objectives of SIMON-I was not only to determine students' RIASEC interest profiles, but at the same time to link this to a set of study programs with matching interest codes. In the absence of an existing classification scheme to describe study programs in terms of their most prevalent RIASEC characteristics, the current project departed from program expert ratings. In support of these ratings, data from six study programs showed that these expert-rated RIASEC program codes demonstrated good congruence with the average interest profiles of the students in these study programs, as indicated by significantly higher C-indexes than the theoretical mean. Importantly, a systematic comparison of students' interest profiles across programs showed that, with only one exception, SIMON-I meaningfully differentiates between students in such a way that interest scores mirrored the theoretical position of the programs in the hexagon. Further, the present study also included occupational RIASEC interest codes as an additional benchmark for the proposed study program codes. Using the interest codes of occupations that are closely aligned with study programs, we were able to demonstrate good levels of congruence for the study fields of *Languages*, *Health Care Management and Policy,*

Bioengineering and *Civil Engineering*. This thus indicates that the interest profiles of these study programs, as determined by the experts, showed strong resemblance (in terms of the RIASEC letter code) with the prescribed interest profiles of corresponding occupations, as listed in the O*NET system. There was moderate agreement between expert and occupational codes for *Clinical Psychology*. The lowest congruence was found for *Economy* programs (C=7). This result parallels findings of Harrington, Feller, and O'Shea (1993), who also established low similarity between empirical program codes and occupational codes in an Economics program. This suggests that there may be a discrepancy between the interest profile of students enrolled in Economics programs and that of people who are employed as economist.

In addition to these psychometric evaluations, we were also interested in the way students perceived the interest profiles that were generated by SIMON-I. After all, this kind of interest assessment is primarily a process of self-exploration (Holland, 1997), and the surplus for test-takers is that they are presented with (a) structured feedback on personal motives that otherwise may risk to remain unnoticed (under the form of the RIASEC interest profile), and (b) concrete study advice (i.e., a list of possible study programs that align with their personal interests). Knowing that such information is well-accepted by test-takers is important because this may heighten the chances that the feedback is actually being taken seriously. Our findings showed that the majority of respondents indeed tended to agree with the interest profile (91.2%) and with the corresponding programs (77.4%) they received. These results are even more optimistic compared to recent work by Sverko, Babarovic, and Medugorac (2014) who reported that 56.3% of their university student sample found that the advice generated by their interest instrument described them well and another 37.5% was neutral. Although only a minority of the respondents in the present study did not agree with their feedback reports, further use and analyses of SIMON-I need to address this.

Finally, with SIMON-I we also introduced a new methodology helping students choosing between fields of study at either the academic or the vocational level. For this purpose, an additional interest scale was developed intended to measure what was labelled the 'Academic factor'. The underlying idea was that embarking on an academic track requires sufficient interest in study activities that are typical for all study programs at this level. In support of this new scale, results indeed indicated mean differences in scores on the 'Academic' scale between students enrolled in academic programs and those in vocational programs, confirming that this scale differentiates between students with more or less interests that are closely aligned to the academic track. Moreover, the analyses also showed that this Academic scale is also distinct from the conceptually related Holland Investigative interest scale. Where the Investigative scale measures the interest in a specific category of study fields where the focus lies on the analysis of physical, medical, or (bio)chemical data and processes, the Academic scale taps into preferences for academic study activities, irrespective of a specific field of interest. For example, the item '*engrossing in a certain subject in order to write a research paper*' refers to an academic activity that is important across all academic programs, ranging from Language (Artistic), and Psychology (Social) majors to Bioengineering (Investigative) study programs.

Limitations and Directions for Future Research

A number of limitations of this work should be acknowledged. First, the profiling of study programs needs more attention. As for now, an expert judgment method was used, by which vocational interest model experts generated RIASEC profiles of study programs. Although this judgment method has proven to be a reliable and valid method to describe study programs (Rounds et al., 1999), subject matter experts from all programs can provide supplementary and perhaps more fine-grained information in the future. Likewise, it would be of great value to add an 'incumbent method' (Rounds et al., 1999) to assign Holland codes to

programs. This implies the use of the empirically established scores per program to refine the profiles generated by experts. Some study programs only had moderate agreement with the proposed RIASEC codes. The mechanisms that are accountable for this moderate agreement require further inquiry.

Second, more work is also needed on the Academic scale. Although there are general score differences between students in the academic and the vocational track, it is still necessary to test whether these differences apply to all fields of study. It is not unthinkable that students in specific vocational programs are more 'Academic' than students in particular academic programs. This requires more data from students enrolled in vocational programs.

Finally, continued data gathering and analyses are warranted for the examination of additional psychometric test requirements, such as test-retest reliability and concurrent and predictive validity. For example, convergent validity evidence could be examined through simultaneous assessment of SIMON-I and widely used interest inventories such as the Self-Directed Search (Holland, 1985). This might also shed light on the added value of SIMON-I. In the longer-run, the secondary education samples should be followed to investigate the validity of SIMON-I to predict study program choice and performance results.

Conclusion

SIMON-I circumvents important limitations of previously developed measures. It is a promising tool that encourages the exploration of study options when making a vocational choice, be it academic or more vocationally oriented. It is expected that this careful exploration of options will boost student success and retention and thus facilitate a smooth transition between secondary and higher education.

References

Armstrong, P. I., Allison, W., & Rounds, J. (2008). Development and initial validation of brief public domain RIASEC marker scales. *Journal of Vocational Behavior, 73*(2), 287-299. doi:10.1016/j.jvb.2008.06.003

Beinicke, A., Pässler, K., & Hell, B. (2014). Does gender-specific differential item functioning affect the structure in vocational interest inventories? *International Journal for Educational and Vocational Guidance, 14*, 181-198. doi:10.1007/s10775-013-9254-y

Betz, N. E., & Fitzgerald, L. F. (1987). *The career psychology of woman.* Orlando, FL: Academic Press.

Borgen, F. H., & Seling, M. J. (1978). Uses of discriminant analysis following MANOVA: Multivariate statistics for multivariate purposes. *Journal of Applied Psychology, 63*(6), 689-697.

Brown, D. (2002). *Career choice and development*. San Fransisco: John Wiley & Sons.

Brown, S. D., & Gore, P. A. (1994). An evaluation of interest congruence indices: Distribution characteristics and measurement properties. *Journal of Vocational Behavior, 45*(3), 310-327. doi:10.1006/jvbe.1994.1038

Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika, 57*(4), 469-497. doi:10.1007/bf02294416

Campbell, V. L. (1987). Strong-Campbell Interest Inventory. *Journal of Counseling & Development, 66*(1), 53-56.

Darcy, M. U. A., & Tracey, T. J. G. (2007). Circumplex structure of Holland's RIASEC Interests across gender and time. *Journal of Counseling Psychology, 54*(1), 17-31. doi:10.1037/0022-0167.54.1.17

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

de Koning, B., Loyens, S., Rikers, E., Smeets, G., & van der Molen, H. (2012). Generation Psy: Student characteristics and academic achievement in a three-year problem-based learning bachelor program. *Learning and Individual Differences, 22*(3), 313-323. doi:10.1016/j.lindif.2012.01.003

Dolnicar, S., & Grun, B. (2007). User-friendliness of answer formats-an empirical comparison. *Australasian Journal of Market & Social Research, 15*(1).

Dolnicar, S., Grün, B., & Leisch, F. (2011). Quick, simple and reliable: Forced binary survey questions. *International Journal of Market Research, 53*(2), 231-252. doi:10.2501/IJMR-53-2-231-252

Eggerth, D. E., & Andrew, M. E. (2006). Modifying the C index for use with Holland codes of unequal length. *Journal of Career Assessment, 14*(2), 267-275.

Einarsdóttir, S., & Rounds, J. (2009). Gender bias and construct validity in vocational interest measurement: Differential item functioning in the Strong Interest Inventory. *Journal of Vocational Behavior, 74*(3), 295-307. doi:10.1016/j.jvb.2009.01.003

Einarsdóttir, S., Rounds, J., Ægisdóttir, S., & Gerstein, L. H. (2002). The Structure of Vocational Interests in Iceland: Examining Holland's and Gati's RIASEC Models. *European Journal of Psychological Assessment, 18*(1), 85-95. doi:10.1027//1015-5759.18.1.85

Furnham, A. (2001). Vocational Preference and P–O Fit: Reflections on Holland's Theory of Vocational Choice. *Applied Psychology, 50*(1), 5-29. doi:10.1111/1464-0597.00046

Germeijs, V., & Verschueren, K. (2007). High school students' career decision-making process: Consequences for choice implementation in higher education. *Journal of Vocational Behavior, 70*(2), 223-241. doi:10.1016/j.jvb.2006.10.004

Grassi, M., Luccio, R., & Di Blas, L. (2010). CircE: an R implementation of Browne's circular stochastic process model. *Behavioral Research Methods, 42*(1), 55-73. doi:10.3758/brm.42.1.55

Grotevant, H. D., & Durrett, M. E. (1980). Occupational knowledge and career development in adolescence. *Journal of Vocational Behavior, 17*(2), 171-182. doi:10.1016/0001-8791(80)90002-0

Hansen, J.-I. C., Collins, R. C., Swanson, J. L., & Fouad, N. A. (1993). Gender differences in the structure of interests. *Journal of Vocational Behavior, 42*(2), 200-211. doi:10.1006/jvbe.1993.1014

Harms, P. D., Roberts, B. W., & Winter, D. (2006). Becoming the Harvard man: Person-environment fit, personality development, and academic success. *Personality and Social Psychology Bulletin, 32*(7), 851-865. doi:10.1177/0146167206287720

Harrington, T. F., Feller, R., & O'Shea, A. J. (1993). Four methods to determine RIASEC codes for college majors and a comparison of hit rates. *The Career Development Quarterly, 41*(4), 383-392. doi:10.1002/j.2161-0045.1993.tb00412.x

Hogerheijde, R.P., van Amstel, B., De Fruyt, F., & Mervielde, I. (1995). *Beroepskeuze Zelf-Onderzoek versie 1995. Een gids voor opleidings-en beroepskeuze. Handleiding-Nederlands/Vlaamse editie.* Lisse: Swets, & Zeitlinger.

Holland, J. L. (1985a). *Making vocational choices: A theory of vocational personalities and work environments (2nd ed.).* Englewood Cliffs, NJ: Prentice Hall.

Holland, J. L. (1985b). *The Self-Directed Search: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments.* Lutz, FL: Psychological Assessment Resources Inc.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53-60.

Hubert, L., & Arabie, P. (1987). Evaluating order hypotheses within proximity matrices. *Psychological Bulletin, 102*, 172-178.

Malgwi, C. A., Howe, M. A., & Burnaby, P. A. (2005). Influences on students' choice of college major. *Journal of Education for Business, 80*(5), 275-282.

Ministerie van Onderwijs en Vorming. (2012). Hoger Onderwijs in cijfers -Addendum. Retrieved from http://www.ond.vlaanderen.be/hogeronderwijs/werken/studentadmin/studentengegevens/HogerOnderwijsInCijfers-addendum20122013.pdf

Murtaugh, P. A., Burns, L. D., & Schuster, J. (1999). Predicting the retention of university students. *Research in Higher Education, 40*(3), 355-371. doi:10.1023/A:1018755201899

Nagy, G., Trautwein, U., & Lüdtke, O. (2010). The structure of vocational interests in Germany: Different methodologies, different conclusions. *Journal of Vocational Behavior, 76*(2), 153-169. doi:10.1016/j.jvb.2007.07.002

Nauta, M. M. (2010). The development, evolution, and status of Holland's theory of vocational personalities: Reflections and future directions for counseling psychology. *Journal of Counseling Psychology, 57*(1), 11-22. doi:10.1037/a0018213

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational Interests and Performance: A Quantitative Summary of Over 60 Years of Research. *Perspectives on Psychological Science, 7*(4), 384-403. doi:10.1177/1745691612449021

OECD. (2011). *Education at glance 2011: OECD indicators*: OECD Publishing.

OECD. (2014). *Education at glance 2014: OECD Indicators:* OECD Publishing.

Pässler, K., Beinicke, A., & Hell, B. (2014). Gender-related differential validity and differential prediction in interest inventories. *Journal of Career Assessment, 22*(1), 138-152. doi:10.1177/1069072713492934

Prediger, D. J. (1982). Dimensions underlying Holland's hexagon: Missing link between interests and occupations? *Journal of Vocational Behavior, 21*(3), 259-287. doi:10.1016/0001-8791(82)90036-7

Reardon, R., & Bullock, E. (2004). Holland's Theory and Implications for Academic Advising and Career Counseling. *NACADA Journal, 24*(1-2), 111-123. doi:http://dx.doi.org/10.12930/0271-9517-24.1-2.111

Rounds, J., Smith, T., Hubert, L., Lewis, P., & Rivkin, D. (1999). Development of occupational interest profiles for O* NET. *National Center for O* NET Development, Raleigh, NC*.

Rounds, J., & Tracey, T. J. G. (1993). Prediger's Dimensional Representation of Holland's RIASEC. *Journal of Applied Psychology, 78*(6), 75-890. doi:10.1037/0021-9010.78.6.875

Rounds, J., & Tracey, T. J. G. (1996). Cross-cultural structural equivalence of RIASEC models and measures. *Journal of Counseling Psychology, 43*(3), 310. doi:10.1037/0022-0167.43.3.310

Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika, 58*(2), 159-194.

Smart, J. C. (1997). Academic Subenvironments and Differential Patterns of Self-Perceived Growth During College: A Test of Holland's Theory. *Journal of College Student Development, 38*(1), 68-78.

Spokane, A. R. (1985). A review of research on person-environment congruence in Holland's theory of careers. *Journal of Vocational Behavior, 26*(3), 306-343.

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychological Bulletin, 135*(6), 859-884. doi:10.1037/a0017364.

Sverko, I., Babarovic, T., & Medugorac, V. (2014). Pictorial assessment of interests: Development and evaluation of Pictorial and Descriptive Interest Inventory. *Journal of Vocational Behavior, 84*(3), 356-366. doi:10.1016/j.jvb.2014.02.008

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson Education.

Tinsley, H. E. A. (2000). The Congruence Myth: An Analysis of the Efficacy of the Person–Environment Fit Model. *Journal of Vocational Behavior, 56*(2), 147-179. doi:10.1006/jvbe.1999.1727

Tracey, T. J. G. (1997). RANDALL: A Microsoft FORTRAN program for a randomization test of hypothesized order relations. *Educational and Psychological Measurement*(57), 164-168. doi:10.1177/0013164497057001012

Tracey, T. J. G., & Rounds, J. (1996). The spherical representation of vocational interests. *Journal of Vocational Behavior, 48*(1), 3-41. doi:10.1006/jvbe.1996.0002

Tracey, T. J. G., & Rounds, J. B. (1993). Evaluating Holland's and Gati's vocational-interest models: A structural meta-analysis. *Psychological Bulletin, 113*(2), 229-246. doi:10.1037/0033-2909.113.2.229

UNESCO. (1997). *Operational Manual for ISCED 1997 (International Standard Classification of Education) (first edition)*. Retrieved from Paris: http://www.uis.unesco.org/Library/Documents/isced97-en.pdf

Wille, B., De Fruyt, F., Dingemanse, S. A., & Vergauwe, J. (2015). A closer look at the psychological diversity within Holland interest types: Construct validation of the Career Insight Questionnaire. *Consulting Psychology Journal: Practice and Research, 67*(3), 234.

Wille, B., Tracey, T. J. G., Feys, M., & De Fruyt, F. (2014). A longitudinal and multi-method examination of interest–occupation congruence within and across time. *Journal of Vocational Behavior, 84*(1), 59-73. doi:10.1016/j.jvb.2013.12.001

Wolniak, G. C., & Pascarella, E. T. (2005). The effects of college major and job field congruence on job satisfaction. *Journal of Vocational Behavior, 67*(2), 233-251. doi:10.1016/j.jvb.2004.08.010

Appendix

English translation of the SIMON-I questionnaire

Part 1: Activities
Mark the YES column for activities you enjoy to do or activities you would like to try. Mark the NO column for activities you would not like to do. If you really don't know what the activity implies, skip the item.

	Activity	YES	NO	SCALE
1	Developing electronic systems			R
2	Analysing the grammatical structure of a sentence			I
3	Helping people with speech disorders			S
4	Recruiting a job candidate			E
5	Monitoring the quality standards for food safety and hygiene			C
6	Analysing and interpreting research results			Ac
7	Repairing malfunctioning electrical equipment			R
8	Carrying out laboratorial analyses			I
9	Designing a poster for an exhibition			A
10	Helping others with their personal problems			S
11	Organising a conference			E
12	Preparing financial reports			C
13	Reading English language scientific articles*			Ac
14	Being responsible for the maintenance of IT hardware			R
15	Analysing statistics			I
16	Designing webpages			A
17	Developing council prevention campaigns			S
18	Presenting new policy propositions			E
19	Collecting quantitative and qualitative data			I
20	Engrossing in a certain subject in order to write a research paper			Ac
21	Develop new methods for industrial production			R
22	Treating diseases in animals			I
23	Editing the sound and images for a movie			A
24	Formulating education and training policies			S
25	Drawing up the budgets			C
26	Doing the follow up on building sites			R
27	Analysing x-rays/brain scans			I
28	Fit out a show room			A
29	Sport guidance for children, the elderly, …			S
30	Formulate a theory about the differences between population groups			I
31	Monitor quality standards			C
32	Writing clear and logically structured texts			Ac
33	Maintaining airplanes			R
34	Investigating the impact of historical people			A
35	Composing a work of music			A
36	Providing guidance for victims			S
37	Selling a product or service			E
38	Calculating prices			C
39	Distinguishing main issues from side-issues in a text			Ac
40	Installing and maintaining computer servers			R
41	Designing an advertising folder			A
42	Providing information about the assistance for the poor			S
43	Drawing up an organisational business or policy plan			E
44	Checking bank transactions			C
45	Starting studying without being asked for			Ac
46	Developing windmill parks			R
47	Prove a theorem			I

Nr	Activity	YES	NO	SCALE
48	Analysing text structures			A
49	Giving travel advice			S
50	Negotiating contracts			E
51	Drawing up a contract			C
52	Looking up sources to give an idea a scientific basis			Ac
53	Investigating chromosomal defects			I
54	Writing scenarios			A
55	Holding tests, questionnaires and in-depth interviews			S
56	Screening the administration			C
57	Reading texts that include formulas, calculations and tables			Ac
58	Working on a drilling rig			R
59	Turning an idea into a film			A
60	Giving care to patients			S
61	Restructuring an organisation or company			E
62	Checking the compliance of regulations			C
63	Drawing conclusions from a mathematical table			Ac
64	Excluding alternative explanations through experiments			I
65	Designing the layout of a hospital			A
66	Advising youngsters regarding their vocational choice			S
67	Exploring new economic markets			E
68	Drawing up the annual report			C
69	Detecting mistakes in arguments			Ac
70	Setting up a festival stage			R
71	Developing a new medicine			I
72	Writing a review			A
73	Giving training in communication skills			S
74	Starting up an enterprise			E
75	Investigating a cost structure			C
76	Setting up, carrying out and evaluating an own research project			Ac
77	Creating a technical drawing			R
78	Putting theories in their historical and social context			I
79	Creating an art piece			A
80	Giving health advice			S
81	Giving health and parenting education			E
82	Calculating expenses			C
83	Disassembling electrical appliances			R
84	Comparing cultures			A
85	Guiding minority groups on the job market			S
86	Conducting a meeting			E
87	Drawing up a timetable			C
88	Measuring a lane			R
89	Supporting and following up foster families			S
90	Attracting sponsors			E
91	Standing in front of a classroom			S
92	Leading a team			E
93	Managing a database			C
94	Collecting soil samples			R
95	Beginning a herbarium (a plant collection)			I
96	Counseling underprivileged people			S
97	Formulating a treatment plan			S
98	Studying the physical endurance of athletes			I

Part 2: Occupations

Mark YES for professions you would like to practice or that you would like to try. Mark NO for professions you would not like to do. If you think a little bit, you probably know most professions. If you really don't know what a profession entails, skip the item.

Nr	Occupation	YES	NO	SCALE

1	Industrial designer			R
2	Civil engineer			I
3	Fashion designer			A
4	Policy advisor in political and international relations			E
5	Recruitment and selection advisor			E
6	Damage expert			C
7	Agricultural technician			R
8	Teacher			S
9	Business economist			C
10	Accountant			C
11	Electrical engineer			R
12	Biologist			I
13	Art/music teacher			A
14	Speech therapist			S
15	Bank manager			C
16	Landscape architect			R
17	Physicist			I
18	Editor			A
19	Student counselor			S
20	Tax supervisor			C
21	Neurologist			I
22	Policy advisor art and culture			A
23	Educator			S
24	Marketing manager			E
25	Safety advisor			C
26	Construction manager			R
27	Historian			I
28	Director			A
29	Communication manager			E
30	Manager (of a company)			E
31	Judge			C
32	Forester			R
33	Researcher			I
34	Graphic designer			A
35	Psychologist			S
36	Lawyer			E
37	Notary			C
38	Mathematician			I
39	Art historian			A
40	Social worker			S
41	Politician			E
42	Pilot			R
43	Pharmacist			I
44	Linguist			A
45	Divorce mediator			S
46	Journalist			A
47	Structural engineer			R
48	Lab assistant			I
49	Photographer			A
50	Nurse			S
51	Advertising campaign manager			E
52	Chemist			I
53	Tax specialist			C
54	Architect			R
55	Artist			A
56	Educational scientist			S
57	Librarian			A
58	Philosopher			I
59	Representative			E

60	Geneticist			I
61	Interior designer			A
62	Estate agent			E
63	Physiotherapist			S
64	Meteorologist			I
65	Sales manager			E
66	Statistician			I

* item specific for students from non-English speaking countries

Chapter 4: Basic mathematics test predicts statistics achievement and overall first year academic success[3]

Abstract

In the psychology and educational science programs at Ghent University, only 36.1 % of the new incoming students in 2011 and 2012 passed all exams. Despite availability of information, many students underestimate the scientific character of social science programs. Statistics courses are a major obstacle in this matter. Not all enrolling students master the basic mathematical skills needed to pass statistics courses. Therefore, we propose a test that measures these skills. Our aim is to examine the predictive validity of the test with regard to the statistics course and also as to overall academic success. The results indicate that a test of very basic mathematics skills helps to identify at-risk students at and before the start of the academic year. The practical implications of these results are discussed. The test aids the efficient use of means for remedial interventions and supports future students in choosing a higher education program that suits their potential.

Introduction

Success rates in higher education are low. The Organization for Economic Co-operation and Development reported that at this level (over 19 OECD countries) 31% of tertiary education students failed to complete a program (OECD and Indicators, 2008). Moreover, the costs of student dropout are high. For example, the Flemish governments' average annual cost per student is $16,000 (Cantillon et al., 2005). Therefore, governments, institutions, and students are in search of factors that can help determine whether someone will pass or not. Attrition problems are manifest worldwide, but the open access policy of the Flanders educational system poses additional challenges. Therefore, before discussing the determinants for academic achievement in the literature, a short framing of the specific Flemish educational context is in order.

The Flanders Education System: Structure and Admission

There are four types of secondary education (SE) programs in Flanders. The first, general SE (GSE), has an emphasis on broad general education and provides a solid foundation for higher education. Second, technical SE (TSE) emphasizes general and technical matters and prepares for a profession or to still pass on to higher education, which is less frequent. Third, secondary arts education (ASE) combines a broad general education with active arts practice. Finally, vocational SE (VSE) is a practice-oriented education in which young people learn a specific profession (Education in Flanders, 2008).

Flemish higher education could be described as binary (Arum et al., 2007). It consists of two main types of programs: academic and professional. Academic programs are mainly organized by universities, whereas university colleges provide professional programs with an emphasis on executive skills. The professional programs lead to a bachelor degree and correspond to the Bologna first cycle programs of 180 European Credit Transfer and Accumulation System (ECTS) (The Bologna Declaration, 1999). Academic programs also lead

to a bachelor degree at first, but the finality is to complement this degree by a master. Academic programs thus correspond to the Bologna two-cycle programs (for a detailed description of the higher education system in Flanders, we refer to Kelchtermans and Verboven (2008)).

Although some SE programs do not typically prepare for higher education, students can enter almost all tertiary education programs when they obtained a degree from any of these four SE programs. There is no numerus clausus, there are no requirements on the grades obtained during SE for access to higher education, and there are, with the exception of medical and artistic programs, no entrance exams. Moreover, there is a policy of high subsidies and very low tuition fees (Kelchtermans and Verboven, 2008), which are typically less than $800/year.

These measures aim to guarantee socially fair access and improve participation in higher education but have the disadvantage that the first year of university is typically a "selection year." In general, after 1 year of studying, not even half of the newly enrolled students pass (Rombaut, 2006). As mentioned, this implies a high cost for students, parents, institutions, and the government (Declercq and Verboven, 2010).

Educational Background and Student Success

Because the system is open to anyone who has completed SE, virtually all programs show a large heterogeneity in SE backgrounds of new incoming students. For example, the amount of mathematics instruction in SE varies between 0 (in VSE programs) and 8h (in some GSE programs) per week. This heterogeneity is reflected in the differences in passing rates, especially in academic bachelor programs which have a focus on research and scientific skills and knowledge. Students with a VSE degree are consistently less successful than those with a general degree, with success rates of students with technical and arts degrees fluctuating between these extremes (Declercq and Verboven, 2010; Rombaut, 2006; Ministerie van Onderwijs en Vorming, 2009; Netwerk studie- en trajectbegeleiding Universiteit Gent, 2012). Even within the group of students with a general secondary degree, there are major differences

in higher education success rates. Students with a general degree focusing on classical languages, mathematics, and science tend to outperform students with a general degree that focuses on modern languages or social sciences.

The average success rate of newly enrolled students at the Faculty of Psychology and Educational Sciences of Ghent University (during the academic years 2005–2006 to 2008–2009) is 49.5% (Netwerk studie- en trajectbegeleiding Universiteit Gent, 2012). During the academic years 2011–2012 and 2012–2013, only 36.1% of these new students passed all exams successfully. So, rates are dropping.

One of the contributing factors to these low success rates is a suboptimal knowledge of what academic programs entail. Many students seem to underestimate the scientific character of programs in the social sciences. Especially statistics courses are a major obstacle in this matter. For example, Murtonen and Lehtinen (2003a, b) showed that social science students rated statistics courses as the most abstract and difficult subject. Many of these students felt that they were non-mathematical persons and as such could not learn mathematical subjects. Some students were even convinced that no relevant information in human sciences can be obtained through quantitative methods.

On the other hand, educational background does not explain completely why some students pass and others do not. For example, a lot of students succeed despite the fact that they come from SE programs that do not prepare them specifically for (academic) tertiary education. This might be the result of the fact that not only cognitive factors contribute to the choice of SE schooling, but also social class (Werfhorst et al., 2003). Hence, the obtained secondary degree does not always reflect the ability of students to cope with the requirements of academic programs in general and the statistics courses specifically. So, there is a clear need not only in students, but also in student counselors and educators for information about students' initial competences and chances of success.

Statistics Courses in the Social Sciences

Most graduate students enrolled in social and behavioral sciences programs worldwide are required to take at least one statistics course and/or a quantitative-based research methodology course as part of their program (Onwuegbuzie, 2003). It is widely acknowledged that statistics and quantitative methods courses cause problems (Murtonen and Lehtinen, 2003a, b), especially for students in social sciences, who generally have less interest and schooling in mathematical subjects. As a result, factors related to success in statistics courses have been the subject of research.

As Lalonde and Gardner stated (1993), most of the variables that have been examined regarding the acquisition of statistical knowledge fall within three broad categories: anxiety, attitudes, and ability.

Several scholars have addressed the influence of attitudes toward statistics (Budé et al., 2007; Cashin and Elmore, 2005; Chiesi and Primi, 2010; Gal and Ginsburg, 1994; Schau et al., 1995). The general conclusion of these studies is that more positive attitudes relate to better exam results (Vanhoof et al., 2006). The negative impact of statistics anxiety on performance has also been widely documented (e.g., Chiesi and Primi, 2010; Macher et al., 2011; Mellanby and Zimdars, 2010; Musch and Bröder, 1999; Vigil-Colet et al., 2008).

In this study, we will focus on the third category: ability. Both very specific abilities, such as spatial visualization ability (Elmore and Vasu, 1980) and general abilities, such as intelligence, have been examined in relation to statistics achievement. Numerous studies have demonstrated a strong positive correlation between general intelligence and educational outcomes (Kuncel et al., 2004). In the current study, the primary concern was that seeing the influx of students with very dissimilar backgrounds, not all enrolling students master the basic mathematical skills needed to pass statistics courses and perhaps also to pass many other

courses that rely on empirical evidence and research. Mathematic ability is therefore the primary variable of interest in the current study.

A few earlier studies have already addressed the importance of mathematical skills for achievement in statistics courses (Chiesi and Primi, 2010; Harlow et al., 2002; Lalonde and Gardner, 1993; Schutz et al., 1998). Garfield and Ahlgren (1988) pointed out that one of the reasons that students have difficulties grasping the fundamental ideas of probability is the fact that many students have underlying difficulties with rational number concepts and basic concepts involving fractions, decimals, and percentages.

Mathematical skills have often been operationalized by previous mathematical achievement (e.g., Musch and Bröder, 1999; Onwuegbuzie, 2003; Tremblay et al., 2000; Wisenbaker et al., 2000). Others have constructed tests to measure mathematical skills (Harlow et al., 2002; Lalonde and Gardner, 1993; Schutz et al., 1998). More recently, Galli et al. (2011) and Johnson and Kuennen (2006) have created a specific test measuring basic mathematics skills. Both studies provide evidence of the significant contribution of these skills to predict results on statistics exams. Galli et al. (2011) found that students with low mathematical ability had significantly lower grades than students with a medium-high ability. Johnson and Kuennen (2006) found that students who answered all basic mathematics questions correctly were likely to earn a half to a full letter grade higher in an introductory business statistics course. Consequently, they raised the question whether basic math skills may be more important than previously recognized. Ballard and Johnson (2004) came to a similar conclusion with regard to an introductory microeconomics course. They found mastery of extremely basic quantitative skills to be the most important factor for course success, even more than American College Testing (ACT) math scores.

None of these studies, however, examined the extent to which these measures discriminate between students passing their first year successfully and those who did not. In

social sciences programs at Ghent University, passing the first year is closely associated with passing the statistics exam: 85.3% of the students that did not pass the first year also failed the statistics exam. Since passing the statistics course is required to pass the first year, none of the students failing the statistics course passed the first year. Of the students that pass the statistics course, 79.2% passes the first year. Seventeen percent of the failing students pass all courses except the statistics course (there are 12 courses in the standard package of 60 ECTS credits). In other settings as well, students have been reported to view statistics courses as a major threat to the attainment of a degree (Onwuegbuzie, 1995). For many students at least, this seems to be not far from the truth.

From Success in the Statistics Course to Overall Academic Success

Observing the generally acknowledged relation between performance in statistics courses and general academic achievement in social science programs, it is surprising that studies examining this relation are, to our knowledge, non-existent. Math subscales of standardized tests (e.g., Scholastic Aptitude Test and ACT) link mathematical ability to academic achievement, but they might lack the specificity to assess mathematical ability necessary for statistics courses in non-mathematical majors (Galli et al., 2011). If basic mathematical skills contribute to the variance in statistics achievement and if there is a high correlation between statistics achievement and general achievement, the question rises whether a basic mathematics test can contribute to the prediction of general academic achievement.

Our aim was to propose and validate an easy-to-administer test that measures basic mathematical skills considered vital to successfully take on an introductory statistics course in an academic bachelor program. This test was therefore not primarily aimed to discriminate between the better performing students. Because of the heterogeneity of new incoming students and the lack of standardized testing in the Flemish education system, this test could especially help identify at-risk students. In addition, we examined to what extent basic mathematical skills

predict overall academic success. To further substantiate this, we examined the relation between the mathematics test and success in non-mathematical courses. As such, this test may offer a valuable tool in the choice of a major in tertiary education.

To summarize, our goal was twofold:

1. Determining whether a basic mathematics test can predict academic achievement in statistics over and above SE background. Can we predict who will pass the statistics exam?

2. Determining whether a basic mathematics test can predict general academic achievement over and above SE background. Can we predict who will pass the first year successfully? To further substantiate this, analyses of success in non-mathematical courses are added.

Method

Instruments

Construction of the mathematics test: construct definition and item generation. To construct the mathematics test, two matters were considered: the mathematical skills that students are supposed to have acquired by the end of SE as described by the Department of Education in Flanders ("Vakgebonden eindtermen derde graad secundair onderwijs-ASO") and the mathematical prerequisites for the introductory statistics course in the bachelors of psychology and educational sciences. The latter were evaluated by teachers and experts in the field of statistics and by faculty guidance counselors who had been administering informal tests of basic mathematical skills to first year students since 2 years.

A pool of items was developed reflecting basic numerical mastery to be achieved after SE and reflecting prerequisites to enroll the introductory statistics course. These items can be subdivided in seven mathematical topics: numerical knowledge and the order of operations, operations with decimal numbers, operations with brackets, operations with fractions, algebra: working with unknown variables, percentages/proportions, and the rule of three. One example question is "If a runner runs on average 1 km in 5 min, how many has he run after 2 h?" (see

Appendix for full 20-item scale). Question format was varied. Open questions, yes/no items, and multiple-choice questions were alternated.

Reliability analysis of the currently studied sample showed a Cronbach's alpha of .76, which shows that scores on the mathematics test were fairly reliable (Field, 2009). This coefficient is acceptable according to recommendations set forth for preliminary and basic research, and it is in line with the mean .77 alpha reported in previous studies (Peterson, 1994).

To avoid cheating, the test was constructed in four different versions in which the sequence of items was varied. To guarantee comparability, the effect of item sequence was checked. An analysis of variance test shows that test scores did not differ across test versions ($F(3,1935)=0.06$, $p=.98$). Hence, all versions were aggregated for further analysis.

Background variables. Information on the background variable SE diploma and number of hours of mathematics instruction in SE was obtained from the university database. Students were asked to give this information when enrolling for the first time at Ghent University.

Achievement measures. The academic year in Flanders starts at the end of September and consists of two semesters. At the end of each semester, exams are organized that cover the courses taken during the past semester. This gives students a first chance to prove that they have acquired the contents of each course. In Flanders, grades in higher education vary between 0 and 20, with a score of 10 as the passing criterion. If a student does not obtain a score of 10 or higher for the taken courses, he or she gets a second chance to pass the exam. Thus, students get two attempts at passing each course during one academic year.

Grades were obtained from the university database. "Statistics score" is the grade obtained in the introductory statistics course irrespective of the amount of chances taken on the exam. Statistics achievement was further operationalized as "passing statistics" (a dichotomous variable that indicated whether a student obtained a grade of 10 or higher or not). Results on

two non-mathematical courses were also analyzed. The introductory psychology course ("passing psychology") and the sociology course ("passing sociology") were selected because of their inclusion in both the psychology and the educational science program, and because these are not methodological or statistical and are introductory courses in the field of study that students signed up for. The contents should therefore be closely aligned to the students' interests.

General achievement was operationalized as "general success rate (GSR)" which is the ratio of the number of credits that a student obtained over the number of credits that he or she subscribed for. Thus, a GSR of 100 means that the student passed all enrolled courses. This rate was further dichotomized as "passing the first year" (yes or no).

Data Collection

The paper-and-pencil test was administered in the second week of the academic year during the introductory statistics class. The advantages of this early administration were threefold. First, in the second week of the academic year, dropout was non-existent or at least very low. Secondly, class attendance decreases as the semester advances (Van Blerkom, 1992). Thirdly, as the semester advances, students gain knowledge and skills that might bias our measures of initial competence and, therefore, confound predictive validity. Thus, assessments early in the semester positively impacted the response rate, and results on the mathematics test were less contaminated by skills and knowledge gained throughout the academic year. All students attending the class were asked to fill out the test, and they were informed that results would be used only for research purposes.

Participants

In the academic years 2011–2012 and 2012–2013, 1,278 new students enrolled at the Faculty of Psychology and Educational Sciences. Of these students, 80.9% filled out the mathematics test, so responses of 1,034 students were analyzed. Eighty seven point two percent

of the sample were females. The proportion of female students is traditionally high in these majors, but this sample proportion was slightly higher than the proportion of female first-generation students (84.6% in the academic years 2011–2012 and 2012–2013). Ninety seven point three percent of the sample was enrolled in the standard package of 60 ECTS credits.

Procedure

First, several t-tests were carried out to determine whether the samples 2011–2012 and 2012–2013 differed significantly with regard to the dependent and independent measures. Next, we examined whether there was a significant relation between the test score and the outcome variables (passing statistics, passing psychology, passing sociology, and passing the first year) through correlational analysis and t-tests. Third, we conducted a preliminary analysis to determine whether the main outcome variables differed as a function of several background variables. If so, these variables were included in logistic regression analysis to determine whether our mathematics test could improve prediction of outcome above and beyond these background variables. Finally, sequential logistic regression was used to determine whether our mathematics test helped in the prediction of the outcome variables.

Results

Cohort Comparison

To determine whether the cohorts of 2011–2012 and 2012–2013 differed significantly on the dependent and independent variables, independent t-tests were carried out.

The t-tests showed that there were no significant differences in passing the first year between students from the 2011–2012 cohort ($M=.42$, SD$=.49$) and those from the 2012–2013 cohort ($M=.38$, $SD=.49$) ($t(1,027)=1.28$, $p=.20$). There were no significant differences in passing the statistics course ($t(1,025)=-.68$, $p=.50$), in passing psychology ($t(1,024)=.01$, $p=1$), or in passing sociology ($t(1,011)=1.66$, $p=.10$) between students from the 2011–2012 cohort ($M=.48$, $SD=.50$; $M=.77$, $SD=.42$ and $M=.68$, $SD=.47$, respectively) and those from the 2012–

2013 cohort (*M*=.50, *SD*=.50; *M*=.77, *SD*=.42 and *M*=.63, *SD*=.48, respectively). Cohort 2011–2012 (*M*=3.54, *SD*= 1.68) and cohort 2012–2013 (*M*=3.60, *SD*=1.50) did not differ with regard to the hours of mathematics instruction in SE (*t*(878)=−.53, *p*=.60). Mathematics test score differences were also insignificant (*t*(1,032)=.77, *p*=.44) between cohort 2011–2012 (*M*=14.85, *SD*=3.55) and cohort 2012–2013 (*M*=14.68, *SD*=3.30).

These tests indicated that there were no significant differences in outcomes, features, or test responses between cohorts. Thus, we felt safe to aggregate the data for further analyses.

Descriptive Statistics and Bivariate Correlations

Descriptive statistics of the mathematics test scores. Scores on the mathematics test varied between 2 and 20 with a mean score of 14.77 (*SD*=3.44). Skewness was −.57 (*SD*=.08). Taking the rule of thumb that a skewness value more than twice its standard error indicates a departure from symmetry (De Laurentis et al., 2010), scores on the mathematics test were negatively skewed. This might indicate a ceiling effect, resulting in worse discrimination at the high end of the scale, which is plausible as the test aimed only to assess basic starting-level competences. Kurtosis (−.16) was less than twice its standard error (.15), indicating that the test scores did not significantly differ from mesokurtic distribution.

Descriptive statistics and correlational analysis. Table 1 provides descriptive statistics of other variables measured and included in the analysis. For all continuous variables (mathematics test score, statistics score, GSR, and hours of mathematics in SE), Pearson correlations are given. Point biserial correlations (rpb) are shown for the dichotomous outcome variables. The statistics course was passed by 48.8% of the sample, the psychology course by 76.7%, the sociology course by 64.2%, and the first year successfully by 40%.

Table 1 Descriptive statistics and correlations

Variable	Mean	SD	Math test score	Statistics score	GSR	Hours of math in SE	Passing statistics	Passing psychology	Passing sociology	Passing the first year
Math test score	14.77	3.44	1.00	.43**	.35**	.37**	.39** r_{pb}	.26** r_{pb}	.25** r_{pb}	.33** r_{pb}
Statistics score	7.70	5.10		1.00	.75**	.32**	.81** r_{pb}	.55** r_{pb}	.53** r_{pb}	.71** r_{pb}
GSR	71.42	33.18			1.00	.23**	.73** r_{pb}	.76** r_{pb}	.71** r_{pb}	.71** r_{pb}
Hours of math instruction in SE	3.44	1.72				1.00	.28** r_{pb}	.19** r_{pb}	.11** r_{pb}	.26** r_{pb}

GSR general success rate, *SE* secondary education, r_{pb} point biserial correlation

** Correlation is significant at the 0.01 level (two-tailed)

Table 2 Differences in passing (statistics course) between SE diploma categories

a. Differences in passing statistics course between SE diploma categories

	GSE1 , GSE2, GSE3	GSE4	TSE, VSE, ASE	Non-Flemish
% Passing statistics per SE category	66.2% a	41% b	10% c	25% b, c
% Failing statistics per SE category	33.8% a	59% b	90% c	75% b, c
N	405	530	60	16

b. Differences in passing the first year between SE diploma categories

	GSE1 , GSE2, GSE3	GSE4	TSE, VSE, ASE	Non-Flemish
% Passing per SE category	57% d	31.9% e	4.8% f	12.5% e, f
% Failing per SE category	43% d	68.1% e	95.2% f	87.5% e, f
N	405	527	62	16

Each letter denotes a subset of SE categories whose column proportions do not differ significantly from each other at the .05 level

An independent sample t-test showed that students passing the statistics exam scored significantly higher on the mathematics test (*M*=16.15, *SD*=2.96) than students not passing the statistics exam (*M*=13.44, *SD*=3.35) (*t*(1,025)=−13.75, *p*<.01). Significant differences in math test score were also found between students passing (*M*=15.25, *SD*=3.23) and failing (*M*=13.11, *SD*=3.64) the psychology exam (*t*(1,024)=−8.66, *p*<.01) and students passing (*M*=15.39, *SD*=3.16) and failing (*M*=13.60, *SD*=3.68) the sociology exam (*t*(1,011)=−7.73, *p*<.01). Students passing the first year successfully also scored significantly higher on the mathematics test (*M*=16.16, *SD*=2.92) than those who did not (*M*=13.86, *SD*=3.45) (*t*(1,027)=−11.55, *p*<.01). Figure 1 shows passed-failed distribution as a function of mathematics test score.
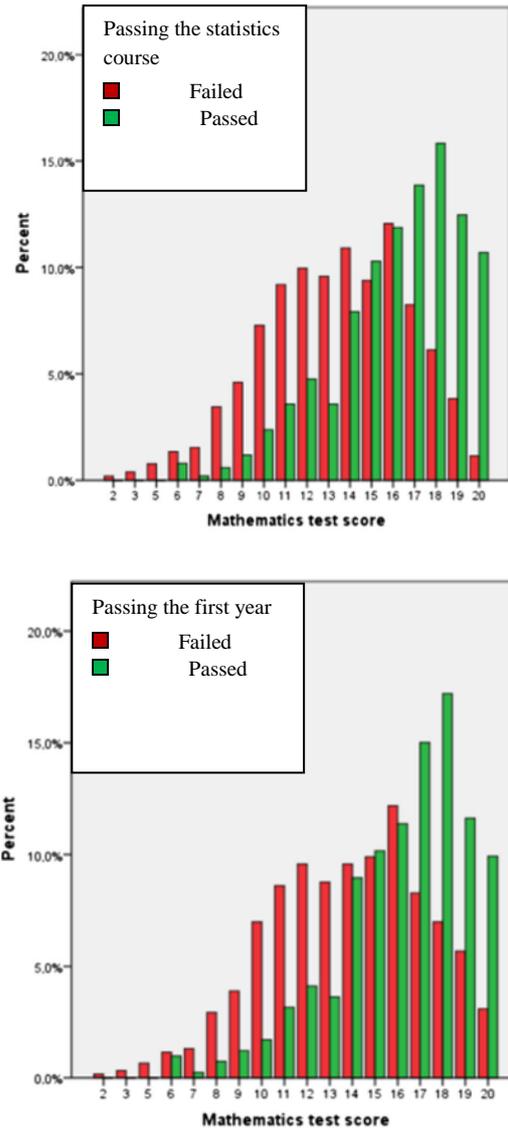

Fig. 1. Distribution of mathematics test scores and passed-failed categories of the statistics course (top) and the first year (bottom).

Since both mathematics test score and hours of mathematics instruction in SE correlated with passing the statistics course and passing the first year, partial correlations were computed between mathematics test score and these outcome variables, controlling for hours of mathematics instruction. The results suggested that the mathematics test score was related to both passing the statistics course (r=.31, p<.01) and passing the first year (r=.24, p<.01).

Moreover, analysis using Steiger's Z tests revealed a stronger relation between mathematics test score and achievement measures (.43, p<.01 for statistics score and .35, p<.01 for GSR) than between the number of hours of mathematics instruction in SE and achievement measures (.32 and .23 respectively, p<.01) (Z=2.61, p<.01 for passing statistics and Z=2.71, p<.01 for passing the first year). This was promising, as it might suggest that the mathematics test is more predictive for achievement than educational background.

Preliminary analysis of background variables

Preliminary analyses were conducted to determine whether the outcome variables passing statistics and passing the first year differed as a function of background variables. If so, the relevant background variables were included in regression analysis as a control.

Gender. To examine the relation between gender and passing the statistics course and the relation between gender and passing the first year, chi-squared tests of independence were performed. There was no significant relation between gender and passing statistics ($\chi2$ (1, N=1027)=.20 p=.65) (47.3 % of the males and 49.4 % of the females passed the statistics exam) or between gender and passing the first year ($\chi2$ (1, N=1029)=2.30, p=.13) (34.1 % of the males and 41 % of the females passed the first year). Gender was not included in the regression analysis.

Educational background: high school diploma. A chi-squared test indicated a significant relation between SE diploma and passing statistics ($\chi2$ (7, N=1008)=131.53, p<.01) and between SE diploma and passing the first year ($\chi2$ (7, N=1010)=130.47, p<.01). To

determine the differences, post hoc Tukey tests were used in ANOVA analysis. Pass-fail distributions as a function of SE diploma are displayed in Figure 2.
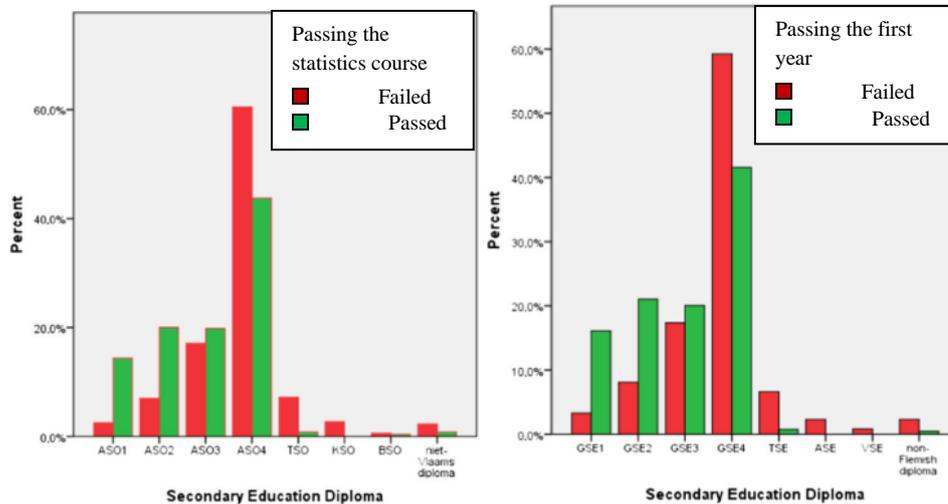
CATEGORIZATION OF SECONDARY EDUCATIONAL BACKGROUND (Rombaut, 2006)
GSE1 (8.1% of sample): General Secondary Education with emphasis on Greek-Latin, Greek-Sciences, Greek-Mathematics, Latin-Mathematics
GSE2 (13.1% of sample): General Secondary Education with emphasis on Latin-Sciences, Mathematics-Sciences
GSE3 (18% of sample): General Secondary Education with emphasis on Latin-Modern languages, Modern languages-Mathematics, Economics-Mathematics
GSE4 (51.3% of sample): General Secondary Education with emphasis on Economics-Modern languages, Modern languages-Sciences, Sport-Sciences, Social sciences
TSE (4.2% of sample): Technical Secondary Education
ASE (1.4% of sample): Artistic Secondary Education
VSE (0.6% of sample): Vocational Secondary Education
Non-Flemish SE (1.5% of sample)

Fig. 2. Distribution of secondary education diploma and passed-failed categories of the statistics course (left) and the first year (right).

Students with a general educational background (GSE1, GSE2, and GSE3) passed the statistics course and the first year more often than students with technical or arts SE. Within the pool of students with a general educational background, students coming from programs with a higher focus on exact sciences and classical languages (GSE1 and GSE2) scored higher than students coming from programs that focus on social sciences, modern languages, and economics (GSE4). The latter did not significantly differ from students with a vocational secondary background. They did differ from students with a technical and arts background on passing the statistics course and from students with a TSE on passing the first year.

None of the students with a background of VSE (*N*=6) passed (the statistics course). Due to this very small cell frequency, the analyses showed no significant differences in passing (statistics) from any other group of students. Students with a non-Flemish diploma (*N*=16) only differed significantly from students with a GSE1 and GSE2 background on both outcomes. High school diploma was considered a relevant variable, but small cell counts obliged us to aggregate data in four categories: a first category consisting of students with a GSE1, GSE2, or GSE3 diploma; a second group of students having a GSE4 diploma; and students with ASE, VSE, and TSE backgrounds were aggregated in group 3 since these students all come from SE programs that do not specifically prepare for higher education studies. Students with a non-Flemish diploma were in group 4. This group was discarded in regression analyses because of the low cell frequencies and because this group was probably very heterogeneous in content. An overview of aggregated group outcomes is presented in Table 2.

Educational background: hours of mathematics instruction in SE. Independent samples t-tests confirmed the positive effect of a stronger focus on mathematics instruction in SE on passing the statistics exam and passing the first year of university (see Figure 3).
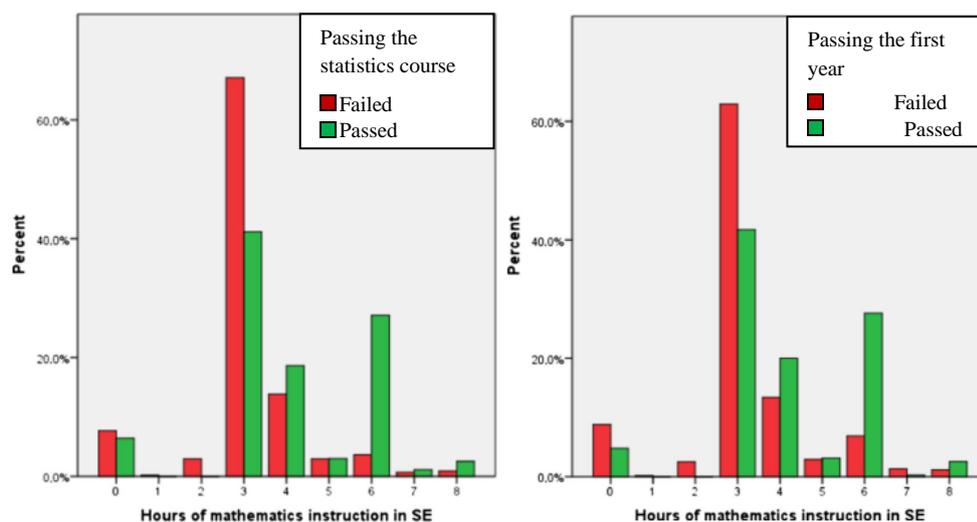
Fig. 3. Hours of mathematics instruction in secondary education and pass-fail distributions of the statistics course (left) and pass-fail distributions of the first year (right).

Students passing the statistics exam had more hours of mathematics in their educational background (M=4.04, SD=1.75) than those who did not pass (M=3.11, SD=1.28) (t(874)= −8.92), p<.01). Students passing the first year successfully also had significantly more mathematics in their SE programs (M=4.08, SD=1.66) than those who did not pass the first year (M=3.22, SD=1.47) (t(876)=−7.98, p<.01).

Hours of mathematics instruction during educational training was hence included in the regression analyses.

Predicting Achievement from Mathematics Score

The next step was determining to what extent our mathematics test could predict achievement. To this end, sequential logistic regressions were conducted using SPSS 19.0.

Predicting statistics achievement from mathematics score. To predict statistics achievement, we ran sequential logistic regressions with passing statistics as dependent variable.

Independent variables entered the regression in two blocks. First, educational background was entered consisting of "SE diploma" category and "hours of mathematics in SE." Secondly, "mathematics test score" was entered.

The chi-squared test statistic of the model with educational background variables alone was statistically significant, $\chi 2$ (3, N=863)=114.89, p<.01. After addition of the mathematics test score, $\chi 2$ (4, N=863)=175.80, p<.01. The total number of correct classifications was 68.5%, which was substantially higher than classification based on the proportion of students failing the statistics course in the sample (48.8%). Comparison of log-likelihood ratios for models with and without the mathematics test score showed significant improvement with the addition of mathematics test score, $\chi 2$ (1863)=60.91, p<.01. Hours of mathematics in SE and mathematics test score did not interact.

A model in which only the mathematics test score was included predicted 65.3% of the cases successfully with Nagelkerke R^2=.21.

The model with all three predictors was best in terms of percentage of correct classifications. Table 3 presents an overview of parameter estimates of this model.

Table 3 Logistics regression parameter estimates and model evaluation – predicting passing statistics course

Secondary education diploma	B	S.E.	Odds ratio
GSE1, GSE2, GSE3 (reference cat.)			
GSE4	-.42***	.17	.66
TSE, ASE, VSE	-1.89***	.58	.152
Mathematics test score	.20***	.03	1.22
Hours of mathematics instruction SE	.22***	.06	1.24
Model evaluation			
Chi-square	175.80***		
Nagelkerke R^2	.25		
Percentage of correct classifications	68.5		

***p<0.001

The score on our mathematics test significantly contributed to the explanation of the variance in passing the statistics exam. The test score alone added 8%, over and above SE background. SE category and the mathematics test score together explained 25% of the variance in passing the statistics exam.

Predicting non-mathematical course results and first year achievement from mathematics score. Sequential logistic regressions were repeated with passing psychology and passing sociology as dependent variables. The mathematics test score added 3.6% of the explained variance of passing psychology on top of the 10.2% explained by educational background factors, $\chi2$ (4, N=864)=80.43, p<.01. For passing sociology, the mathematics test score added a significant 2.9% of the explained variance over and above the 8.4% explained by

background variables, χ^2 (4, N=855)=72.89, p<.01. These results showed that the mathematics test score aids the prediction of success in non-mathematical courses as well.

To predict general achievement, passing the first year was used as dependent variable. A model in which only the mathematics test score was included predicted 67.1% of the cases successfully with Nagelkerke R^2=.15. Analysis showed that a model with educational background variables alone improved prediction of passing the first year, χ^2 (3, N=865)=109.25, p<.01. After addition of the mathematics test score, χ^2 (4, N=865)=140.08, p<.01, Nagelkerke R^2=.20. The percentage of correct classifications was 67.5%, which was substantially higher than classification based on the proportion of students failing the first year in the sample (40%). Comparison of the models proved a significant contribution of the mathematics test score to the prediction of passing the first year successfully, χ^2 (1, 865)=30.83, p<.01. This was the best model in terms of percentage correct classifications. The mathematics test score added 4.2% to the explained variance in passing the first year, over and above SE background variables. Table 4 presents an overview of parameter estimates of this model.

Table 4 Logistics regression parameter estimates and model evaluation – predicting passing first year successfully

Secondary education diploma	B	S.E.	Odds ratio
GSE1, GSE2, GSE3 (reference cat.)			
GSE4	-.53**	.17	.59
TSE, ASE, VSE	-3.20**	1.03	.04
Mathematics test score	.14***	.03	1.15
Hours of mathematics instruction SE	.18***	.06	1.20
Model evaluation			
Chi-square	140.08***		
Nagelkerke R^2	.20		
Percentage of correct classifications	67.5		

** p<0.01; ***p<0.001

The score on the mathematics test significantly contributed to the explanation of the variance in passing the first year successfully. The mathematics test proved to have predictive validity over and above SE background. Twenty percent of the variance in passing the first year successfully was explained by SE background and the mathematics test score, with the mathematics test score adding 4.2% of the explained variance in passing the first year. The model accurately predicted 67.5% of the cases.

Discussion

Statistics and methodology courses are often seen as a threat to the attainment of a degree (Onwuegbuzie, 1995), especially in social science programs. Success rates in the current study confirm this perception. Of the students who passed the statistics course, 79.2% passed the first year successfully, and none of the students failing the statistics course passed their first year.

Despite the relatively low performance on statistics examinations, only a few studies have yet investigated predictors of success in such courses (Budé et al., 2007; Galli et al., 2011; Kennett et al., 2009; Macher et al., 2011; Onwuegbuzie, 2003). Basic mathematical skills have been found a relevant variable in the prediction of statistics achievement, a variable that may be more important than previously recognized (Johnson and Kuennen, 2006).

If basic mathematical skills contribute to the variance in statistics achievement and there is a high correlation between statistics achievement and general achievement, the question rises whether a basic mathematics test can contribute to the prediction of general academic achievement. In this study, we answered this question positively. To our knowledge, this study was the first to address the relation between basic mathematical skills needed to succeed in introductory statistics courses and overall academic success.

We constructed a test that is short and easy to administer, assessing extremely basic mathematical skills considered vital to pass introductory statistics courses. Results showed that

the mathematics test score significantly contributed to the prediction of passing the statistics course, adding 8% of the explained variance over and above SE diploma and hours of mathematics instruction in SE. Moreover, the mathematics test score also explained variance in the non-mathematical courses of introductory psychology and sociology. This supported us in the proposition that basic mathematics skills contribute to the prediction of overall academic achievement. The basic mathematical test score did indeed explain, together with SE diploma, 20% of the variance in passing the first year successfully or not. The mathematics test score alone accounted for 4.2% of the explained variance. The model correctly classified 67.5% of the cases.

This number is not sufficient to identify all at-risk students and to properly inform students on expected capacities in higher education, but it certainly helps in alerting some individuals with potential deficits in the required numerical mastery. It is noteworthy that a test of very basic mathematical problems significantly contributes to the prediction of overall academic success. Moreover, the test is very brief, consisting of only 20 items, which makes it quick and easy to administer.

These results have several practical implications. The results confirm the implicit feeling of teaching staff and student counselors that knowledge of basic math operations is vital to academic success in the first year at university and that there is a group of students that lack this basic knowledge. The basic mathematical skills test allows identification of the students that have a high probability of failing the statistics course and, as a consequence, the first year of the psychology or educational science program. After assessment, enrolled students that lack these basic mathematical skills can be encouraged to take up remedial courses in mathematics. Moreover, the test allows pupils, SE teachers, and student counselors to evaluate objectively whether or not someone has acquired the necessary mathematics skills to pass an introductory statistics course. As such, the test is valuable not only for students that are already enrolled in

the programs, but also for pupils that are in the process of choosing a field of study in higher education. The test aids potential students in evaluating whether they master the required skills to pass the introductory statistics course and their first year of higher education. This is priceless information in a system that has open access to higher education, that has no standardized tests, and where not all pupils have mathematics in their SE curriculum.

This study also shed light on the relation between SE background and academic achievement in higher education social science programs. We found that students with more hours of mathematics instruction in SE had significant better results in higher education. Regardless of the hours of mathematics instruction, the specific SE program also showed related to both statistics achievement and overall first year academic achievement. Overall, students from GSE programs passed significantly more often than did students from technical, arts, or VSE programs. This is not surprising, since the latter programs focus more on direct entrance into the labor market whereas only the GSE programs specifically prepare for continuation into higher education. Nevertheless, even within GSE programs, we found significant differences in success rates. Students coming from programs with a higher focus on exact sciences and classical languages performed better than students coming from programs that focus on social sciences, modern languages, and economics. Although these results are in line with previous data (Declercq and Verboven, 2010; Rombaut, 2006), they are somewhat surprising. One would expect that a background in economics, which relies on mathematical principles, would provide a sufficient basis for statistical courses in higher education. Even more deterrent is the fact that students coming from a SE program that focuses on social sciences have relatively low success rates in tertiary education programs in that same field of study. One possible explanation could be that students from "more difficult" SE programs have a heavier workload which encourages the development of specific study skills that allow them to cope more effectively with higher education demands. A second reason could be that SE social sciences programs might focus

less on the methodological and statistical components that are crucial in the higher education curricula of these majors. As such, wrong ideas of what these programs entail could be fostered. The explanation that is most often proclaimed is the selection effect (Declercq and Verboven, 2010), whereby "weaker" students in the course of their SE gradually choose programs that are viewed as "easier," such as the social sciences program. Nevertheless, these underlying reasons have as yet not been studied and open up an interesting area of research.

There are a few limitations to our study. First, mathematics test scores were negatively skewed. This was an inevitable consequence of the test goal, because the purpose was to detect insufficiencies in basic knowledge. Therefore, skewness in test score distribution does indicate a ceiling effect. This ceiling effect may limit the size of the observed correlations in the current study, because the discriminatory power of the test is aimed at the lower end of the scale. The present correlations may hence reflect a lower bound estimate of true predictive validity of post-SE numerical tests. By adding extra items with higher difficulty, the test might be able to discriminate more at the high end. Second, we only addressed the ability to predict first year grades from the mathematics test. We cannot yet determine to what extent the test has predictive value for long-term results such as persistence and timely graduation. A more longitudinal approach is thus recommended and is on our agenda. Next, it would be interesting to examine whether the relation between the mathematics test and passing is found in other fields of study, for example, in programs where statistical courses are not compulsory. Finally, only basic mathematical skills and SE background were taken into account. There are many additional individual differences that have been studied in relation to academic achievement, such as intelligence (e.g., Busato et al., 2000; Kuncel et al., 2004), personality traits (e.g., De Fruyt and Mervielde, 1996; Lounsbury et al., 2003; O'Connor and Paunonen, 2007), and motivational factors (Budé et al., 2007; Steinmayr and Spinath, 2009). Likewise, psychosocial and study skills factors have been observed to significantly contribute to the prediction of academic

achievement (Robbins et al., 2004). Despite the overwhelming amount of literature on predicting academic achievement, the combination of cognitive and non-cognitive factors is less often studied.

The mathematics test is only a beginning in the search for a model that takes into account most relevant factors in the prediction of academic success. It is our intention to add different variables to the model to increase predictive power.

More background variables, other skills and personality, and motivational and self-efficacy factors will be added in future work. By combining cognitive and non-cognitive factors, we hope to develop a model that provides increased predictive validity and is thus valuable to educators, student counselors, and students. Our general aim is to help identifying at-risk students and to help future students in evaluating their cognitive and non-cognitive abilities in order to choose a major that best suits their potential and background.

References

Arum, R., Gamoran, A., & Shavit, Y. (2007). *More inclusion than diversion: expansion, differentiation, and market structure in higher education.* In Shavit, R. & Gamoran, A. (Eds.), Stratification in higher education: a comparative study (pp. 1–35). Palo Alto: Stanford University Press.

Ballard, C. L., & Johnson, M. F. (2004). Basic math skills and performance in an introductory economics class. *Journal of Economic Education, 35(1),* 3–23. doi:10.3200/jece.35.1.3-23.

Budé, L., Van De Wiel, M., Imbos, T., Candel, M., Broers, N., & Berger, M. (2007). Students' achievement in a statistics course in relation to motivational aspects and study behaviour. *Statistics Education Research Journal, 6(1)*, 5–21.

Busato, V., Prins, F., Elshout, 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*, 1057–1068.

Cantillon, B., Verbist, G., & Segal, I. (2005). *Student in de 21ste eeuw. Studiefinanciering voor het hoger onderwijs in Vlaanderen*. Rapport voor Onderwijskundig Beleids-en Praktijkgericht Wetenschappelijk Onderzoek (OBPWO): Ministerie van de Vlaamse Gemeenschap, Departement Onderwijs.

Cashin, S., & Elmore, P. (2005). The survey of attitudes toward statistics scale: a construct validity study. *Educational and Psychological Measurement, 65(3)*, 509–524. doi:10.1177/0013164404272488.

Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal, 9(1),* 6–26.

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

De Laurentis, G., Maino, R., & Molteni, L. (2010). *Developing, validating and using internal ratings: methodologies and case studies*. West Sussex: Wiley Ltd.

Declercq, K., & Verboven, F. (2010). Slaagkansen aan vlaamse universiteiten: tijd om het beleid bij te sturen? *Impuls, 42(2),* 88–98.

Education in Flanders. *A broad view of the Flemish educational landscape*. (2008). Retrieved 4 April 2013, from http://www.eui.eu/Documents/MWP/AcademicCareers/Countries/Belgium/BelgiumFlemishHigherEducation.pdf

Elmore, P. B., & Vasu, E. S. (1980). Relationship between selected variables and statistics achievement: building a theoretical model. *Journal of Education & Psychology, 72(4)*, 457–467. doi:10.1037/0022-0663.72.4.457.

Field, A. (2009). *Discovering statistics using SPSS*. London: Sage Publications Ltd.

Gal, I., & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: towards an assessment framework. *Journal of Statistics Education, 2(2)*.

Galli, S., Chiesi, F., & Primi, C. (2011). Measuring mathematical ability needed for "non-mathematical" majors: the construction of a scale applying IRT and differential item functioning across educational contexts. *Learning and Individual Differences, 21(4),* 392–402. doi:10.1016/j.lindif.2011.04.005.

Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: implications for research. *Journal for Research in Mathematics Education, 19(1)*, 44–63.

Harlow, L., Burkholder, G., & Morrow, J. (2002). Evaluating attitudes, skill, and performance in a learning enhanced quantitative methods course: a structural modeling approach. *Structural Equation Modeling: A Multidisciplinary Journal, 9(3)*, 413–430. doi:10.1207/s15328007sem0903_6.

Johnson, M., & Kuennen, E. (2006). Basic math skills and performance in an introductory statistics course. *Journal of Statistics Education, 14(2)*.

Kelchtermans, S., & Verboven, F. (2008). Participation and study decisions in a public system of higher education. *Journal of Applied Econometrics, 25(3)*, 335–391.

Kennett, D., Young, A. M., & Catanzaro, M. (2009). Variables contributing to academic success in an intermediate statistics course: the importance of learned resourcefulness. *Educational Psychologist, 29(7)*, 815–830. doi:10.1080/01443410903305401.

Kuncel, N., Hezlett, S., & Ones, D. (2004). Academic performance, career potential, creativity, and job performance: can one construct predict them all? *Journal of Personality and Social Psychology, 86(1)*, 148–161. doi:10.1037/0022-3514.86.1.148.

Lalonde, R., & Gardner, R. (1993). Statistics as a second language? A model for predicting performance in psychology students. *Canadian Journal of Behavioural Science, 25(1)*, 108–125. doi:10.1037/h0078792.

Lounsbury, J., Sundstrom, E., Loveland, J., & Gibson, L. (2003). Intelligence, "big five" personality traits, and work drive as predictors of course grade. *Personality and Individual Differences, 35*, 1231–1239. doi:10.1016/s1475-1585(03)00053-5.

Macher, D., Paechter, M., Papousek, I., & Ruggeri, K. (2011). Statistics anxiety, trait anxiety, learning behavior, and academic performance. *European Journal of Psychology of Education, 27(4)*, 483–498. doi:10.1007/s10212-011-0090-5.

Mellanby, J., & Zimdars, A. (2010). Trait anxiety and final degree performance at the University of Oxford. *Higher Education, 61(4),* 357–370. doi:10.1007/s10734-010-9335-5.

Ministerie van Onderwijs en Vorming. (2009). Studiesucces Generatiestudenten in 2007–2008.

Murtonen, M., & Lehtinen, E. (2003a). Difficulties experienced by education and sociology students in quantitative methods courses. *Studies in Higher Education, 28(2)*, 171–186.

Murtonen, M., & Lehtinen, E. (2003b). Difficulties experienced by education and sociology students in quantitative methods courses. *Studies in Higher Education, 28(2)*, 171–185. doi:10.1080/0307507032000058064.

Musch, J., & Bröder, A. (1999). Test anxiety versus academic skills: a comparison of two alternative models for predicting performance in a statistics exam. *The British Journal of Educational Psychology, 69*, 105–116.

Netwerk studie- en trajectbegeleiding Universiteit Gent. (2012). *Studievoortgang bij de generatiestudenten aan de Universiteit Gent*.

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

OECD, & Indicators. (2008). *Education at a glance 2008*. London: Routledge.

Onwuegbuzie, A. (1995). The effect of time constraints and statistics test anxiety on test performance in a statistics. *Journal Of Experimental Education, 63(2)*, 115.

Onwuegbuzie, A. (2003). Modeling statistics achievement among graduate students. *Educational and Psychological Measurement, 63(6)*, 1020–1038. doi:10.1177/0013164402250989.

Peterson, Robert A. (1994). A meta-analysis of Cronbach's coefficient alpha. *Journal of consumer research*, 381–391

Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130(2)*, 261–288. doi:10.1037/0033-2909.130.2.261.

Rombaut, K. (2006). *Determinanten van differentiële slaagkansen in het hoger onderwijs.* Universiteit Antwerpen, Centrum voor sociaal beleid Herman Deleeck: Onuitgegeven onderzoeksrapport.

Schau, C., Stevens, J., Dauphinee Thomas, L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement, 55(5)*, 868–875.

Schutz Paul, A., Drogosz Lisa, M., White Vicki, E., & Distefano, C. (1998). Prior knowledge, attitude, and strategy use in an introduction to statistics course. *Learning and Individual Differences, 10(4)*, 291–308. doi: 10.1016/S1041-6080(99)80124-1.

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences, 19(1)*, 80–90. doi:10.1016/j.lindif.2008.05.004.

The Bologna Declaration of 19 June 1999. Joint declaration of the European Ministers of Education. (1999).

Tremblay, P., Gardner, R., & Heipel, G. (2000). A model of the relationships among measures of affect, aptitude, and performance in introductory statistics. *Canadian Journal of Behavioural Science, 32(1)*, 40–48. doi:10.1037/h0087099.

Vakgebonden eindtermen derde graad secundair onderwijs-ASO) http://www.ond.vlaanderen.be/curriculum/secundair-onderwijs/derde-graad/aso/vakgebonden/wiskunde/eindtermen.htm. Accessed 5 Apr 2013.

Van Blerkom, M. (1992). Class attendance in undergraduate courses. *The Journal of Psychology, 126(5)*, 487-494.

Vanhoof, S., Sotos Ana Elisa, C., Onghena, P., Verschaffel, L., Van Dooren,W., & Van den Noortgate,W. (2006). Attitudes toward statistics and their relationship with short-and long-term exam results. *Journal of Statistics Education, 14(3)*, 3.

Vigil-Colet, A., Lorenzo-Seva, U., & Condon, L. (2008). Development and validation of the statistical anxiety scale. *Psicothema, 20(1)*, 174–180.

Werfhorst, H. G., Sullivan, A., & Cheung, S. Y. (2003). Social class, ability and choice of subject in secondary and tertiary education in Britain. *British Educational Research Journal, 29(1)*, 41–62. doi:10.1080/0141192032000057366.

Wisenbaker, J., Scott, J., & Nasser, F. (2000). Structural equation models relating attitudes about and achievement in introductory statistics courses: a comparison of results from U.S and Israel. Tokyo: *Paper presented at the Ninth International Congress on Mathematics Education.*

Appendix

Twenty-item mathematics test

Item	Question	Answer format
1	You write a number with the digit 1, the digit 2, and the digit 3. All of these digits are used precisely one time. How many different numbers can you write?	Open answer
2	If a runner runs on average 1 km in 5 min, how many has he run after 2 h?	Open answer
3	Complete: If $x/y=0.25$, then $y/x=$	Open answer
4	Calculate:$8-0.8*5=$	Open answer
5	Calculate: A book has a 40 % discount and costs €18. What was the price of the book before the discount was subtracted?	Open answer
6	Calculate:$4*0.8=$	Open answer
7	Complete: If $x-y=0.5$, then $y-x=$	Open answer
8	Is this expression correct? $(a+b)+(-c-d)=(a-c)-(b-d)$	a. Yes b. No
9	Calculate:$3/5*(-2)+1/5=$	Open answer
10	In a group of 400 people, there are 270 men and 130 women. The proportion of low-skilled women is 0.4. How many women in this group are low-skilled?	Open answer
11	Calculate: $(-1)-(-6)$	Open answer
12	What is smaller than 1?	a. $1/2+5/9$ b. $7/8+1/4$ c. $2/3+4/12$ d. $2/5+1/4$ e. None of the above
13	Calculate: The square root of 0.01	Open answer
14	What is correct?	a. $0.023>0.05$ b. $0.05>0.023$
15	$f(x)$ is the amount of gas in my car in function of the distance x (in km, since the last time I filled up the car). $f(x)=50-0.05x$. How many kilometers can I drive before the tank is empty?	Open answer
16	Find x. $2x2+4=3x2-5$ $x=$	Open answer
17	Calculate: $2/16$	Open answer in percentage
18	A car uses 6 L of gas in 100 km. How much gas does it use in 250 km?	Open answer
19	Calculate: $2/3*3/2=$	Open answer
20	Complete: $24=75$ % of	a. 0.32 b. 18 c. 32 d. 36 e. None of the above

Chapter 5: Program-Specific Prediction of Academic Achievement on the Basis of Cognitive and Non-cognitive Factors[4]

Abstract

Choosing a suitable study program is one of the factors that facilitates academic achievement and thus prevents drop-out in the first year of tertiary education. This requires adequate information on both the individual abilities and the environment during the study choice process. The SIMON (Study Skills and Interest MONitor) project of Ghent University, Belgium, provides this information to prospective students through an online tool that informs them a) on the match between their interests and study programs and b) about their personal chances of success in specific study programs. The current study intends to validate the prediction of program-specific chances of success by examining a) the (incremental) predictive validity of cognitive and non-cognitive variables of conscientiousness, motivation, self-efficacy, metacognition and test anxiety and b) the differential predictive power of variables within and across study programs. In addition, a path model with structural relations between variables was tested. The sample consisted of 2391 new incoming students.

Results supported the incremental validity of non-cognitive factors. Achievement could be predicted by cognitive and background factors and by conscientiousness, self-efficacy and test anxiety. Moreover, the predictive power of variables varied across study programs, which suggests that research findings about the prediction of academic achievement might benefit from taking into account the specific program context.

Practical implications for research and (educational program choice) counselling of students are discussed.

Study context: Flanders and the SIMON Project

Drop-out rates in higher education are high. The Organisation for Economic Co-operation and Development reported that 32% of incoming tertiary students do not graduate from a program at this level (OECD, 2008). Vocational choice, and more specifically choice of program of study or major, is certainly an important topic in this matter. According to person-environment fit theories, choosing an educational program that fits the individual is one of the factors that facilitates academic success and can thus prevent drop-out in the first year of tertiary education. For example, the Minnesota Theory of Work Adjustment posits that a person's achievement and satisfaction is predicted from the correspondence between the abilities of the person and the ability requirements of the environment (Dawis, 2005). In order to make an optimal study choice, adolescents should identify their values and abilities, as well as the educational possibilities that correspond with these values and abilities (Swanson & Schneider, 2013). This requires adequate information on both the individual and the environment during the study choice process. When potential students are able to assess their personal abilities and their fit with educational programs, this may increase student retention (McGrath et al., 2014). Moreover, providing an instrument that assesses these factors may increase social equality in higher education as it are often socially vulnerable groups that lack the information to make a realistic educational program choice or to enroll in tertiary education (Müller, 2014; OECD, 2003).

Although universally relevant, such an assessment tool is especially valuable in the current study context, Flanders, which is the northern region of Belgium. Flanders has a public education system where access to higher education is almost unconstrained. The majority of higher education systems across the world use some form of examination (e.g. the Scholastic Aptitude Test in the U.S.) or rely on a minimal secondary education academic performance in the admission process. In Flanders, however, admission restrictions virtually don't exist. Any

student with a secondary education qualification can enter almost any higher education institution and field of study. With the exception of medicine, dentistry and performing arts programs, there are no selection exams, there are no entrance quota and secondary education Grade Point Average (GPA) is never considered for admission. On top of that, tuition fees are extremely low (below $1000 per year). This system is assumed to foster social mobility and to improve participation of economically disadvantaged groups in higher education, but the open entrance implies de facto that the first year of university is typically a "selection year". Less than 40% of university students pass all courses during the first year of studying (even after repeated examination attempts). This is alarming, especially because first year performance is one of the best predictors of academic retention (de Koning et al., 2012; Murtaugh et al., 1999).

In addition to open access, students must enroll in a specific study program and select a major already at the start of higher education. Therefore, in the current paper the term 'study program' refers to both the choice of program of study and of the specific major. Switching programs usually requires students to restart as a freshman. Taken together, the study options are numerous and (financial and motivational) consequences of selecting an inappropriate program are high. This context makes the study orientation process even more important and the provision of adequate information on the match between a prospective student and a specific study program even more crucial.

In response to these challenges, Ghent University started the SIMON-project (Study skills and Interest MONitor), developing a freely available online assessment tool by which students can assess their interests (SIMON-I, Fonteyne, Wille, Duyck, & De Fruyt, 2016) and competencies (SIMON-C). As admission is free by law, SIMON is not an admission tool, but it is designed to provide prospective students (*before* enrollment) with relevant information on the match between their interests/competencies and study programs and on program-specific chances of success in tertiary education. The assumption is that adequate and personalized

information will help students to make better higher education study choices. As stated by McGrath et al. (2014), this can be achieved by introducing non-selective entry tests and strengthening pre-university orientation, which is exactly the objective of the SIMON-project.

The focus in the current study is on the evaluation of competencies with regards to specific study programs (SIMON-C). As such, its purpose is to identify whether prospective students have low chances of success in specific study programs, based on historic data of students with comparable abilities. In contrast with high-stake admission tests, SIMON-C's discriminatory power lies at the lower end of the ability range: its aim is to identify a small group of students that has a very low probability of passing. This is also in accordance with the open access policy: only potential students who almost certainly lack the very basic abilities to succeed (should) get a clear warning, yet, students who may be vulnerable but who might still be able to pass get the benefit of doubt and are not be discouraged. In short, SIMON-C targets to predict tertiary academic achievement (and especially failure) relying on the student's skills and abilities. Assessment of skills and abilities in SIMON-C was based on the vast amount of studies pertaining to the prediction of academic success and retention.

What Factors Predict Academic Achievement?

Cognitive factors.

The use of cognitive ability to predict academic success has a long standing tradition. In fact, the first broad test of cognitive ability (the Binet-Simon scale in 1905) was specifically designed to predict achievement in an educational context. Since then, cognitive ability, or g (a construct related to intelligence) has been consistently found to predict academic achievement (Ackerman & Heggestad, 1997; Busato, Prins, Elshout, & Hamaker, 2000; Farsides & Woodfield, 2003; Kuncel, Hezlett, & Ones, 2004). As the importance of cognitive ability for academic achievement has been well documented, a detailed overview is beyond the scope of this study. It suffices to say that many authors argue that cognitive ability is (one of) the

strongest predictor(s) of academic performance (Kuncel & Hezlett, 2010; Petrides, Chamorro-Premuzic, Frederickson, & Furnham, 2005), with correlations with GPA ranging from .30 to .70 (Roth et al., 2015). As a result, it is mainly cognitive ability that is tested for admission decisions in countries with restricted access to higher education. Most of these tests assess a combination of verbal and quantitative skills (Sedlacek, 2011).

In many predictive studies of academic achievement, previous academic achievement (often high-school GPA) is also taken into account. However, high-school GPA has the great disadvantage that it is not comparable across high schools (and even teachers). Moreover, studies indicate that grades have become a less useful indicator of student success, mainly because of "grade inflation" (Sedlacek, 2011). Therefore, in the current study we included hours of mathematics instruction in secondary education as a background factor, as previous data and research have shown that this is a relevant predictor in the current study context (Fonteyne et al., 2015). Note that Flanders does not have a common, standardized exam (like the SAT) at the end of secondary education.

Non-cognitive factors.

Although cognitive factors are highly relevant in the prediction of academic achievement, correlations between ability measures and academic performance are lower at more advanced levels of education (Boekaerts, 1995), which is generally explained by range restriction effects (e.g., Furnham & Chamorro-Premuzic, 2004; Richardson, Abraham, & Bond, 2012; Sternberg, Grigorenko, & Bundy, 2001). Also, some students fail in spite of high cognitive ability and some students compensate a lack of cognitive or test-taking ability by showing greater motivation or effective study strategies (Komarraju, Ramsey, & Rinella, 2013). Therefore, assessment of other factors is also valuable.

Allen, Robbins, and Sawyer (2009, p.2) define non-cognitive factors as "nontraditional predictors that represent behavioral, attitudinal, and personality constructs, primarily derived

from psychological theories". 'Non-cognitive' refers to a variety of constructs. As a result, several classifications have been proposed. De Raad and Schouwenburg (1996) noted that Messick (1979) provided an encompassing list of potential non-cognitive factors, which included background factors, attitudes, interests, temperament, coping strategies, cognitive styles, and values. Lipnevich and Roberts (2012) proposed a taxonomy of four categories: attitudes and beliefs (self-efficacy), social and emotional qualities, learning processes and personality. Sedlacek (2010) mentioned, apart from others, positive self-concept, realistic self-appraisal and also the ability to handle racism. This shows that the classification of these constructs is not straightforward which prompts a selection of relevant predictors depending on the context.

Apart from cognitive factors, personality has been proposed as one of the main determinants of academic achievement arguing that cognitive factors would measure maximal performance (what can the student do?) whereas personality would account for typical performance (what will the student do?) (Chamorro-Premuzic, Furnham, & Ackerman, 2006). Indeed, many studies have shown that (Big Five) personality factors add incremental predictive validity for academic achievement over and above cognitive factors (see e.g., Poropat, 2009). Especially *Conscientiousness* has been raised as an important predictor for academic success (Conard, 2005; Noftle & Robins, 2007; Poropat, 2009; Trapmann, Hell, Hirn, & Schuler, 2007). Therefore, conscientiousness was included in the current study.

As for other non-cognitive constructs, we chose to include only factors for which predictive validity for academic achievement has been demonstrated over and above cognitive factors. This allowed to limit testing time and was in accordance with our aim to advise prospective students based on a scientifically valid tool. We turned to meta-analyses to identify such non-cognitive constructs as these summarize the results of multiple studies and therefore generate more robust estimates of reliable effect sizes. We came across two large meta-analyses

that fit our purposes. They are both well cited and examined the effect of non-cognitive constructs over and above cognitive predictors.

A first is a study by Robbins et al. (2004), which included 109 studies. They found that the best non-cognitive predictors of college GPA were academic self-efficacy and academic motivation (ρs .50 and .30, respectively). Academic self-efficacy was a better predictor than both high school GPA and ACT/SAT scores (ρs .45 and .39, respectively). A second meta-analysis (by Credé & Kuncel, 2008) examined the incremental validity of study skills, habits and attitudes such as self-regulatory skills and time management. They found that study motivation and study skills exhibit the strongest relationship with GPA (ρs .39 and .33, respectively). Academic-specific anxiety was an important negative predictor of performance (ρ = -.18). Based on these studies, we chose to include these relevant variables in our research.

Robbins et al. (2004) identified academic self-efficacy as an important predictor. Self-efficacy (Bandura, 1997) is described as 'beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments'. Choi (2005) and Pajares (1996) demonstrated that particularized measures of self-efficacy have better results in the prediction and explanation of related outcomes. As a consequence, academic self-efficacy has been empirically related to academic achievement (Bong, 2001; Chemers, Hu, & Garcia, 2001; Choi, 2005; Elias & Loomis, 2002; Galyon, Blondin, Yaw, Nalls, & Williams, 2012; Lent, Brown, & Larkin, 1986; Multon, Brown, & Lent, 1991; Owen & Froman, 1988; Vuong, Brown-Welty, & Tracz, 2010; Zajacova, Lynch, & Espenshade, 2005). Ferla (2008) found that academic self-efficacy explained 7.4% of the variance in Psychology and Educational Sciences students' academic performance. Therefore, academic self-efficacy could not be lacking in the current study. Still, some have argued that high self-efficacy has detrimental effects. For example, Vancouver, Thompson, Tischner, and Putka (2002) have found a negative relationship between self-efficacy and performance. This might be a result of high self-efficacy leading to diminished

effort which in turn affects performance negatively (Vancouver & Kendall, 2006). As a result, it may be reasonable to distinguish several dimensions of self-efficacy, which we took up in the current study. We examine two dimensions of self-efficacy: one called 'effort' (the confidence one has that one will put in the effort to succeed) and another labeled 'comprehension' (the confidence one has that one will understand the contents of the courses). Whilst the first is expected to have a positive relation with academic achievement, the latter may indicate an overestimation of one's abilities which leads to a decrease in effort and results in lower performance.

Study motivation predicted academic achievement in both meta-analyses. We used motivation from a self-determination perspective (SDT). In SDT, motivation is multidimensional in that it distinguishes two qualities of motivation (Vansteenkiste, Sierens, Soenens, Luyckx, & Lens, 2009): autonomous and controlled. Autonomous motivation involves engaging in an activity out of personal interest or relevance. In contrast, controlled motivation involves doing a task with a sense of pressure or guilt. Deci and Ryan (2000) noted a convergence between SDT and Achievement Goal Theories (e.g., Dweck, 1986). According to these authors, autonomous motivation is practically equivalent to learning goals. Yet, they also state that performance goals do not align well with the construct of controlled motivation because performance goals can also be pursued for autonomous reasons. We follow their argumentation that it is necessary to not only consider *what* goals people chase (e.g., performance goals), but also *why* they pursue them (for autonomous or controlled reasons). Therefore, we included motivation from an SDT perspective. Motivation has been shown to impact academic performance, with positive effects for especially autonomous motivation (Bailey & Phillips, 2015; Kusurkar, Ten Cate, Vos, Westers, & Croiset, 2013; Taylor et al., 2014). For example, Vansteenkiste, Zhou, Lens, and Soenens (2005) found that autonomous motivation accounted for 6% of the variance in exam performance of Chinese students.

Credé and Kuncel (2008) emphasized self-regulatory skills as important factors, which refer to the processes which maintain the cognition, affect, and behavior necessary to achieve intended goals (Schunk & Zimmerman, 1997). Cognitive components of self-regulation such as metacognition have been studied. Spada and Moneta (2014) showed that maladaptive metacognition promotes a surface approach to learning ($r = .42$) which in turn leads to poor academic performance ($r = -.33$). González and Paoloni (2015) found that autonomy support, motivation and the metacognitive strategies of planning, monitoring and evaluation explained 57% of the variance of the grade in a chemistry course.

Effects of test anxiety, which was also identified in the meta-analysis of Credé and Kuncel (2008), on performance have been somewhat mixed. De Raad and Schouwenburg (1996) stated that correlations between test anxiety and academic performance are generally low, with typical values between 0.10 and .20. On the other hand, numerous studies have supported that test anxiety does have a detrimental effect on performance (see for example Byron & Khazanchi, 2011; Eysenck, Derakshan, Santos, & Calvo, 2007; Hembree, 1988; Hill & Wigfield, 1984). A possible explanation for these opposing findings is that studies measure different aspects of test anxiety. Liebert and Morris (1967) introduced the idea that test anxiety consists of two components: worry and emotionality. Worry refers to the cognitive concern about test taking and performance, such as negative expectations, preoccupation with performance, and potential consequences. Emotionality refers to perceived physiological reactions, that is, autonomic arousal and somatic reactions to testing situations such as nervousness and tension (Hong & Karstensson, 2002). Since then, this bi-dimensionality has been widely accepted (Cassady & Johnson, 2002) and research has indicated that it is especially the cognitive − or worry − component that (negatively) influences achievement (Kitsantas, Winsler, & Huie, 2008; Morris, Davis, & Hutchings, 1981; Seipp, 1991), which is thus included in the current study.

An integrative model.

The structural relationship between most of these variables has previously been addressed in the control-value theory of achievement emotions (Pekrun, 2006). Of the variables included in the present study, he proposed that achievement emotion (test anxiety) is influenced by motivation and self-efficacy and that all of these, combined with cognitive and metacognitive factors, affect performance. The current study allows us to test this model and to extend it by adding conscientiousness since personality was not included in his control-value theory of achievement emotions.

Combination of predictive factors.

Research has shown that cognitive factors as well as non-cognitive skills predict tertiary academic achievement, yet simultaneous investigations of these conceptually very different factors are scarce, and most studies focus on a specific antecedent of academic success. Also, some studies include variables that are not measurable before, or at the start of tertiary education and therefore did not allow prediction of academic performance *before* enrollment, as is our goal. Exemplary is a Dutch study by de Koning et al. (2012), in a sample of 1753 students, which was nevertheless restricted to a Psychology program. They showed the relative contribution of observed learning activities, first- and second-year performance, high school grades, conscientiousness, and verbal ability towards academic achievement in the bachelor program ($R^2 = .30$). Likewise, Dollinger, Matyja, and Hubert (2008, U.S., $N = 338$), examined verbal ability, the five-factor model, GPA, academic goals, attendance and study behavior to predict academic achievement in a Psychology course ($R^2 = .46$).

Other studies are more in line with our objective to make predictions based on start-of-the-year competencies, yet they do not include all of the variables that are currently under scrutiny. For example, personality factors were not included in a study by Kitsantas et al. (2008). They did however find that 47% of the variance in students' academic achievement was

accounted for by the combination of prior ability levels (cumulative high school GPA and verbal and math SAT scores), self-regulatory processes, and motivational beliefs (U.S. sample, $N = 243$). Personality variables were also missing in a study by Olani (2009), as were metacognitive skills and test anxiety. With regards to other variables, they found that the combination of prior academic achievement (preparatory school GPA, aptitude test scores, and university entrance exam scores) and psychological variables (achievement motivation and academic self-efficacy) accounted for 17% of the variance in students' university GPA scores. The sole contribution of psychological variables was 4 % (Ethiopia; $N = 214$, from departments of Electrical Engineering, International Trade and Investment Management, Information System Management, Mathematics and Psychology).

The Ridgell and Lounsbury (2004) study (U.S., $N = 140$) did include personality, but left out metacognitive skills, test anxiety and motivational factors. General intelligence, (big five) personality traits and work drive explained 24% of the variance of course grade in introductory psychology and of self-reported GPA.

Thus, none of these studies combined cognitive ability measures with personality, motivation, self-efficacy, metacognition and test anxiety in the prediction of academic success. To our knowledge; only one study did include all factors currently under scrutiny. Richardson et al. (2012) conducted a meta-analysis of 55 European and 186 Northern American data sets that included demographic factors, measures of cognitive capacity or prior academic performance and 42 non-cognitive constructs from 5 research domains: (a) personality traits, (b) motivational factors, (c) self-regulatory learning strategies, (d) students' approaches to learning, and (e) psychosocial contextual influences. They found that performance self-efficacy ($r = .59$) was the strongest correlate of GPA, followed by high-school GPA ($r = .40$), ACT ($r = .40$), and grade goal (the GPA the student intends to attain) ($r = .35$).

This study relied mainly on United States samples. As U.S. universities are highly selective, these results may be biased by problems with range restriction, i.e. the predictive value of variables (or their relative importance) may differ in the sample of freshmen that have actually been allowed in tertiary education, relative to the population. Although several methods allow researchers to correct for such bias, these methods are not flawless (Wiberg & Sundström, 2009). The current study is less hindered by restriction of range effects as there are no selection criteria in the higher education system in Flanders, and given that a majority of students also fails the selected program. Moreover, predictions from prior studies do not take into account context-specific differences, such as the specific program for which academic achievement is investigated.

Differential Prediction: Disciplinary Differences

There is an abundance of studies on academic achievement. Surprisingly, the field of study in which the subjects were recruited is rarely a subject of discussion. Still, the few studies that have addressed this issue show evidence for disciplinary differences in the predictive power of variables. For example, Vedel, Thomsen, and Larsen (2015) studied the variability of predictive power of personality traits (both broad and narrow) across academic majors. They found that the R^2 of Big Five personality facets ranged from .16 (Arts/Humanities majors) to .57 (Psychology majors). Vanderstoep, Pintrich, and Fagerlin (1996) studied self-regulated learning and found that the relationships between adaptive motivational beliefs and academic performance differed as a function of academic discipline. Shaw, Kobrin, Patterson, and Mattern (2012) found that the relationship between the SAT and GPA varied by major. The SAT was most predictive of GPA in STEM (science, technology, engineering and mathematics) fields. Also, three sections of the SAT differed in their predictive power. SAT-writing tended to be the strongest predictor for most majors, although SAT-mathematics was the strongest for biological and biomedical sciences (r = 0.59), engineering/architecture (r = 0.57), and

mathematics and statistics/physical sciences (r = 0.59) majors, while SAT-critical reading was the strongest for security and protective services (r = 0.55) majors as well as social services and public administration (r = 0.55) majors.

If there is variablity in predictive power between academic disciplines, this has consequences for both counselling (prospective students should be able to evaluate their competences with regard to specific fields of study) and research practice (findings of differential predictive power might also influence the generalizability of results from studies using specific samples in the prediction of academic success). Especially when students do not know yet in which program they are going to enroll, it would be valuable if a common tool that assesses a broad range of academic/cognitive competences allows program specific predictions. As such, a prospective student could use one broad test to evaluate which programs fit his/her personal profile.

The Present Study

As stated, studies on the combined predictive validity of a wide range of cognitive and non-cognitive factors are scarce. Thus, a first aim of the current study is to examine the incremental predictive validity of cognitive factors, background variables, personality, self-efficacy, motivation, metacognition and test anxiety in the prediction of academic achievement in a large sample of students having open access to higher education. In doing so, we also test a structural model paralleling the control-value theory of achievement emotions (Pekrun, 2006) in which all variables predict achievement and test anxiety is predicted by motivation and self-efficacy.

In addition to the scarcity of studies on the combination of cognitive, non-cognitive and background factors, the differential predictive validity of these variables across different tertiary education programs is rarely examined. It has been argued repeatedly that different fields of study require different competencies (see e.g., Holland, 1997; Stark & Lowther, 1988).

For example, technical fields of study require a higher level of mathematical skills, whereas arts majors require a higher level of verbal skills (Morgan, 1990). Such specificity makes it likely that the predictive power of variables varies across higher education disciplines. If this is the case, prospective students would benefit from the opportunity to evaluate their personal skills with reference to specific fields of study as opposed to receiving generalized feedback on their compentence level. Note that such program-specific prediction is especially challenging on the basis of a common test, for students who are still exploring multiple program options. Surprisingly, few studies have addressed this issue. Yet, findings on differential predictive power of variables across academic disciplines would also have consequences for the generalizability of results from studies using specific samples, and consequently specific academic disciplines, in the prediction of academic success.

In sum, the aim of the current study is twofold:

1. To examine the validity of the combination of background variables, cognitive factors, conscientiousness, metacognition, motivation and test anxiety for the prediction of academic achievement in a sample that is less hindered by restriction of range.

Hypothesis 1

In the current sample that is more heterogeneous than student samples that have been pre-selected to attain a tertiary education program, cognitive and background variables will explain a considerable amount of variance in academic achievement across programs.

Hypothesis 2

Conscientiousness, motivation, metacognition and test anxiety will explain variance in academic achievement over and above background and cognitive factors.

Hypothesis 3

The self-efficacy dimension 'effort' will have a positive relation with academic achievement whereas its dimension 'comprehension' will have a negative relation with outcomes.

Hypothesis 4

In a structural model, all variables will predict academic achievement, and test anxiety will be predicted by motivation, self-efficacy and conscientiousness.

2. To examine variations in predictive power of factors across academic disciplines.

Hypothesis 5

Program-specific predictions will explain more variance in academic achievement and will lead to higher classification success (i.e., will allow a higher percentage of correctly identified at-risk students) than overall-sample predictions.

Because of the exploratory nature of the current study we have few hypothesis regarding the role of non-cognitive factors in the different study program. With regards to the cognitive factors, we do expect the following:

Hypothesis 6

Verbal skills will be more important in Law and Languages programs as a result of their emphasis on languages.

Hypothesis 7

Mathematical skills will be more important in Psychology and Pharmaceutical sciences as these programs include statistical courses.

Method

Procedure

At the start of the academic year, all new incoming undergraduate students across 5 faculties of Ghent University were invited, both orally and by email, to fill out the instrument. Students who had not completed the instrument by the second week of the academic year received a reminder by email. A second reminder was sent at the end of the second week and the assessment was closed down at the beginning of the third week of the academic year. At the end of the academic year, exam results (both binary pass/fail and GPA) were retrieved from the

university database. The study was approved by the faculty ethics committee and students gave written consent for participation.

Participants

Programs were only considered for inclusion in the study if their response rate was higher than 60% and if there was a minimum of 120 respondents (10 respondents per predictor variable). This led to inclusion of 8 programs of 5 faculties. The overall response rate in these programs was 79%, leaving a sample of 2391 subjects for further analysis. 71.7% of these respondents were female, which is marginally lower than the average number of female students enrolling in the included programs, which is 71.9% (Ministerie van Onderwijs en Vorming, 2014, i.e. Department of Education). 35.8% of the sample passed the first year successfully, which does not deviate from general passing rates in the current study context (Ministerie Van Onderwijs en Vorming, 2009). An overview of the included programs and the response rate, gender and passing rate is given in Table 1.

Table 1 Sample Characteristics (Response Rate, Gender and Passing Rate)

Faculty	Program	Response rate	Sample *N*	% Female	% Passed
Psychology and Educational Sciences	Psychology	90.2	744	82.4	39.6
Law	Law	90.3	449	61.5	24.3
	Criminology	74.3	135	72.1	23.9
Arts and Philosophy	Linguistics and literature	69.9	316	74.4	52.6
	History	67.2	172	33.1	21.2
	Applied linguistics	71.8	147	73.5	33.3
Veterinary Medicine	Veterinary medicine	91.6	220	75.5	40.6
Pharmaceutical Sciences	Pharmaceutical sciences	82.6	208	78.8	41.3

Measures

An overview of included variables and descriptive statistics for each study program is given in Table 2.

Background characteristics.

The hours of mathematics instruction that the respondents received in secondary education was retrieved from the university database.

Cognitive factors.

Basic mathematic skills were measured using a 20-item instrument that assesses basic numerical competence, not specific mathematics knowledge. One example item is "Calculate: A book has a 40 % discount and costs €18. What was the price of the book before the discount was subtracted?". Respondents were not allowed to use calculators, but they could use scrap paper to write down calculations. There was no time limit. This instrument has been shown to predict academic achievement in the current study context (Fonteyne et al., 2015). Cronbach α in the current sample was .62. Factorial structure was examined using exploratory factor analysis. With the exception of one, all items loaded on one factor. When examining solutions with more factors, these did not indicate multidimensionality of the scale. Therefore, we decided to use the scale as previously validated.

Reading comprehension consists of an English text with 5 multiple choice questions. This text was previously validated and used in the Swedish Scholastic Aptitude Test. Cronbach α in the current sample was .32. Internal consistency was low because of the limited number of items and because responses were skewed. Items were answered correctly by most respondents. Yet, as it is the purpose to identify students that lack basic competencies, it is valuable to identify which students fail to answer these questions correctly. All items loaded on one factor.

Vocabulary knowledge was administered with the LexTALE (Lemhöfer & Broersma, 2012). Respondents are asked to indicate whether 60 items (e.g., 'pastitie') are existing Dutch

words or not. The resulting percentage score is an indication of general Dutch proficiency (α = .75). Factorial structure was examined. With the exception of five, all items loaded on one factor. When extracting more factors, these five items did not seem to load on a distinct factor. Therefore, we decided to use the scale as previously validated.

Non-cognitive: conscientiousness.

Conscientiousness was measured using the PfPI (De Fruyt & Rolland, 2010), which is a Big Five personality measure that has been validated in the Flemish context. The C-scale consists of 48 items that are rated on a 5-point Likert scale (e.g., I am a well-organized person). Cronbach α was .91. All items loaded on one factor.

Non-cognitive: self-efficacy, motivation, metacognition and test anxiety.

Academic self–efficacy was measured with an adapted version of the College Academic Self-Efficacy Scale by Owen and Froman (1988). Alpha internal consistency estimates of 0.90 and 0.92 are reported, and the stability across an 8-week period was 0.85. As "social" academic aspects such as "talking to a professor privately to get to know him or her" do not, or only to a lesser extent, apply to undergraduate programs at universities in Flanders, these items were excluded from the original scale, resulting in 22 items. Students used a 5-point Likert scale to indicate their self-efficacy levels. Factor analysis showed that the items loaded on two factors, identified as 'effort' (N = 8, loadings between .468 and .736, e.g., "Attending class regularly", α = .76) and 'comprehension' (N = 14, loadings between .416 and .636, e.g., "Understanding most ideas you read in texts", α = .79).

Motivation: A Flemish adaptation (Vansteenkiste, 2009) of the Academic Self–Regulation Questionnaire (Ryan & Connell, (1989) was administered. Respondents indicated on a 5–point Likert scale to what extent they agree with different reasons for studying. Items for controlled (N = 8, e.g., "because I'm supposed to do so", α = .87) and autonomous motivation (N = 8, e.g., "because I want to learn new things", α = .85) were included. Factor

analysis confirmed that items loaded their respective factors (loadings between .692 and .763 for controlled and between .494 and .820 for autonomous motivation).

Metacognition: The Metacognitive Awareness Inventory (Schraw & Dennison, 1994) was used, which has two main subscales: Knowledge of cognition (17 items, $\alpha = .87$, example item "I am good at remembering information") and Regulation of cognition (35 items, $\alpha = .93$, example item "I consider several alternatives to a problem before I answer"). As in De Backer, Van Keer, and Valcke (2012), the original scoring system was replaced with a six-point Likert-type scale ranging from 1 (I totally disagree) to 6 (I totally agree). Factorial structure was partially confirmed. Although loadings were below the threshold of .40 (Stevens, 2012) for 2 items on the Knowledge of cognition scale and for 4 items on the Regulation of cognition scale, all items loaded on the proposed factor.

Cognitive test anxiety was assessed using the Cognitive Test Anxiety Scale Revised (CTAR) (Cassady & Finch, 2015). The CTAR measures the cognitive domain of test anxiety. Participants responded to 25 items such as "While preparing for a test, I often think that I am likely to fail" using a Likert-scale ranging from 1 to 4. The respondent's total score represents the level of cognitive anxiety. Prior reliability analyses have shown high internal consistency. For example, Cassady (2004) found a Cronbach's alpha of 0.93. In this sample, responses indicated the same high reliability ($\alpha = .93$). Unidimensionality of the scale was confirmed with exploratory factor analysis.

Outcome variables

The main dependent variable is whether or not students pass the first year successfully. In Flanders, uniform passing criteria are used across faculties. Students pass the first year when they obtain a credit for all courses taken, which means they scored a minimum of 10 out of 20 on the exam. Moreover, assessment methods are fairly uniform in the first year of higher education. In all included study programs, multiple choice and open answer formats are

standard. In about 10 to 20% of the courses, these written exams are complemented with coursework and participation credits. This standardization of passing criteria and of examination form allows comparison across study programs. Analyses with GPA (max. 1000) as the dependent variable are also included in order to facilitate comparison with international literature, even though SIMON was designed to optimally predict passing rates at the lower end of study success.

Analytic Procedure

As it is our intention to examine whether a specific cluster of variables significantly adds to the model's ability to predict the probability of passing, we used hierarchical logistic regression (Tabachnick & Fidell, 2007). Although the binary outcome is of specific interest to our study and counselling practice with the instrument, GPA is often used in research on academic achievement. In order to allow for comparison, we also performed hierarchical linear regressions with GPA as the dependent variable. Independent variables entered the regressions in four blocks. The order was based on previous research on academic achievement. First, traditional predictors were entered: educational background first because at the point of assessment, this could not be altered. This was followed by a block of cognitive factors. Next, we included conscientiousness, as this has previously been identified as an important personality variable. To assess the incremental validity of other non-cognitive factors, we entered motivation, self-efficacy, metacognition and test anxiety. Given that SIMON-C is constructed to identify those prospective students who have a very low probability of passing, classification success is also evaluated. These regressions are complemented with a path analysis (with maximum likelihood estimation) in which all variables predict GPA and in which motivation, self-efficacy and conscientiousness predicted academic emotion test anxiety.

Results

Preliminary Analysis and Descriptive Results

Table 2 shows the mean scores and standard deviations for all tests and study programs. Zero-order correlations between variables are reported in Table 3. Prior to analyses, multicollinearity was examined and Variance Inflation Factor (VIF) values were all well below 10 (Stevens, 2012). The residuals histogram showed a fairly normal distribution, which indicated that the normality of residuals assumption was satisfied.

The correlations of test scores with the outcome variables passing and GPA (shown in Table 4) confirmed many of the expected relationships. With few exceptions, background and cognitive predictors were significantly related to the outcome variables, as were the non-cognitive predictors conscientiousness, test anxiety and self-efficacy. There was only one faculty (Arts and philosophy) in which all included predictors were significantly related to academic achievement. In the Veterinary medicine students however, only background and non-cognitive predictors (metacognition, motivation and self-efficacy) were associated with achievement. Contrary to all other programs, cognitive ability predictors, conscientiousness and test anxiety failed to reach significance in Veterinary medicine students.

Prediction of Passing

Table 5 shows the (increase in) explained variance for each cluster of variables (ß's are shown in Table 6). A regression analysis on the total sample yielded significant results for all groups of variables. Background and cognitive variables explained respectively 6 and 8% of the variance in passing which confirmed our first hypothesis. Our second hypothesis, that non-cognitive variables would explain variance over and above traditional predictors, was also affirmed. The explained variance (Nagelkerke R^2) was .180. Program-specific analyses generated higher explained variances, varying between .179 and .282 with an average of .233, confirming hypothesis 5.

In seven out of eight study programs, more than one cluster of variables significantly added to the prediction of academic success. In three of the study programs, the motivation and test anxiety cluster was significant (with ΔR^2 between .08 and .18). In Criminology, this was the only significant cluster. Conscientiousness significantly predicted passing over and above background and cognitive factors in four of the eight study programs (with ΔR^2 between .02 and .06).

When looking at the significant contribution ($p < .05$) of specific variables to the prediction of passing (Table 7), the combination of traditional predictors (cognitive and background factors) and non-cognitive predictors (personality, self-efficacy, metacognition, motivation and test anxiety) allowed for the best prediction in 5 out of 8 programs as evidenced by a significant ΔR^2. For Criminology, Applied linguistics and Veterinary medicine, only non-cognitive variables predicted passing. Supporting hypothesis 6, we did find verbal skills (reading comprehension and vocabulary) to be important in the Law and Linguistics and literature programs. Yet, contrary to our expectations, these skills did not contribute to the prediction of passing in Applied linguistics. Also, verbal skills significantly predicted passing in the Psychology program. Hypothesis 7 was also partially confirmed. As expected, mathematical skills were important in Psychology and Pharmaceutical sciences, but they were also significant in the Law, Linguistics and literature and History programs. Hypothesis 3 was confirmed: The self-efficacy dimension 'effort' was only significant in the Psychology program, but it was positively related to passing whereas the 'comprehension' dimension had a negative relation with passing in all 4 programs in which it was significant.

Table 2 Descriptive statistics of included variables for each study program.

	Psychology		Law		Criminology		Linguistics and literature		History		Applied linguistics		Veterinary medicine		Pharmaceutical sciences	
	M	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
Mathematics hours SE	3.85	1.38	4.10	1.46	3.56	1.20	3.65	1.19	3.78	1.52	3.54	1.07	5.12	1.58	5.50	1.39
Mathematics test	15.41	2.89	14.37	2.42	14.30	2.79	15.58	3.12	14.64	3.79	14.66	2.73	15.85	2.48	18.08	1.95
Reading	3.53	1.19	3.91	0.99	3.47	1.21	3.83	1.00	3.67	1.13	3.76	1.19	3.67	1.18	3.76	1.17
Vocabulary knowledge	87.60	8.56	89.27	7.40	87.18	9.42	89.69	7.85	89.87	7.53	87.76	9.42	87.75	9.85	87.98	7.73
Conscientiousness	150.15	18.37	157.82	20.91	147.13	16.86	153.70	19.80	144.22	18.14	145.36	20.68	151.02	19.29	151.50	20.67
Test anxiety	52.53	12.42	48.07	12.02	53.12	11.90	48.67	12.64	51.44	13.30	51.78	12.91	50.89	13.12	53.20	13.21
Metacognition (knowledge)	68.56	10.31	72.70	10.04	66.91	10.39	70.29	10.52	65.07	11.62	67.25	11.32	69.60	10.00	69.84	11.67
Metacognition (regulation)	134.49	20.64	141.82	20.66	134.21	20.33	135.53	21.62	127.18	23.06	131.19	22.82	135.75	22.45	137.89	22.56
Controlled Motivation	16.45	6.09	16.22	5.95	16.11	5.87	17.65	6.58	16.50	6.14	18.35	6.86	14.77	5.60	17.84	7.03
Autonomous Motivation	30.64	4.72	30.85	4.77	29.48	4.67	31.19	4.80	29.79	5.33	29.53	5.29	31.12	4.70	30.58	4.50
Self-Eff. (effort)	29.90	4.00	30.69	3.98	29.43	3.75	30.62	3.67	28.48	4.32	29.69	3.96	30.39	3.59	30.61	3.68
Self-Eff. (comprehension)	49.78	5.53	51.54	5.94	49.04	5.28	50.60	6.01	48.81	6.72	48.81	6.17	50.65	5.50	51.42	5.39

Table 3 Zero-order correlations between predictor variables

	1	2	3	4	5	6	7	8	9	10	11
1. Mathematics hours in SE	-										
2. Mathematics test	.339**	-									
3. Reading Comprehension	.078**	.198**	-								
4. Vocabulary knowledge	-.122**	-.254**	-.155**	-							
5. Conscientiousness	-.007	.023	.038	-.008	-						
6. Test anxiety	-.025	-.076**	-.106**	.026	-.265**	-					
7. Metacognition (knowledge)	.029	.103**	.124**	-.134**	.561**	-.293**	-				
8. Metacognition (regulation)	.027	.092**	.073**	-.165**	.573**	-.123**	.781**	-			
9. Controlled Motivation	-.012	.021	-.024	-.021	-.040	.240**	.012	.046*	-		
10. Autonomous Motivation	-.018	.028	.050*	-.038	.470**	-.092**	.443**	.471**	.022	-	
11. Self-Efficacy (effort)	.013	.049*	.017	-.091**	.592**	-.222**	.488**	.490**	-.036	.453**	-
12. Self-Efficacy (comprehension)	.177**	.226**	.122**	-.095**	.365**	-.348**	.511**	.433**	-.016	.374**	.466**

** $p < .01$ * $p < .05$

Table 4 Correlations of predictor variables with outcomes Passing and GPA per Faculty

Faculty	Psychology and Educational Sciences		Law		Arts and Philosophy		Veterinary medicine		Pharmaceutical Sciences	
	Pass	GPA	Pass	GPA	Pass	GPA	Pass	GPA	Pass	GPA
Mathematics hours in SE	.29**	.29**	.21**	.31**	.12**	.13**	.16*	.17*	.26**	.34**
Mathematics test	.27**	.27**	.21**	.33**	.21**	.29**	.10	.10	.21**	.22**
Reading Comprehension	.22**	.20**	.13**	.14**	.17**	.22**	.07	.11	.16*	.18*
Vocabulary knowledge	.16**	.10**	.12**	.14**	.09*	.14**	.06	.08	.08	.15*
Conscientiousness	.12*	.12**	.04	.09*	.23**	.27**	.07	.07	.24**	.26**
Test anxiety	-.12**	-.12**	-.14**	-.13**	-.22**	-.17**	-.08	-.04	-.12	-.23**
Metacognition (knowledge)	.10*	.10*	.04	.04	.22**	.25**	.11	.12	.08	.06
Metacognition (regulation)	.09*	.08*	.02	.03	.16**	.17**	.14*	.13	.10	.10
Controlled Motivation	.06	.07	-.02	-.01	.10*	.12**	-.01	-.01	-.02	-.02
Autonomous Motivation	.06	.08*	-.00	.03	.15**	.17**	-.14*	-.06	.00	.02
Self-Eff. (effort)	.11**	.13**	.03	.04	.16**	.22**	.16*	.17*	.19**	.24**
Self-Eff. (comprehension)	.07*	.04	.10*	.08	.16**	.14**	-.01	.03	-.01	.03

Note. SE = Secondary Education; Self-Eff. = Self-Efficacy. *$p < .05$ **$p < .01$

Table 5 (increase in) explained variance of Passing for each cluster of variables

Program	N	R^2 background	ΔR^2 Cognitive skills	ΔR^2 Conscientiousness	ΔR^2 Motivation and test anxiety	Total R^2
Psychology	744	.112**	.087**	.019**	.021	.254
Law	449	.099**	.084**	.003	.024	.207
Criminology	135	.000	.036	.006	.176*	.218
Linguistics and literature	316	.012	.079**	.059**	.081**	.231
History	172	.026	.154**	.003	.099	.282
Applied linguistics	147	.055*	.024	.056*	.097	.232
Veterinary medicine	220	.035*	.016	.010	.118**	.179
Pharmaceutical sciences	208	.092**	.080**	.034*	.051	.257
Average		.054	.070	.024	.083	.233
Total	2391	.055**	.077**	.018**	.030**	.180

Note. **p < .01 *p < .05

Table 6 ß Coefficients of Logistic Regressions (Passing)

Program	Hours of math in SE	Mathematics	Reading comprehension	Vocabulary	Conscientiousness	Test anxiety	Metacognition (knowledge)	Metacognition (Regulation)	Controlled motivation	Autonomous motivation	Academic self-efficacy Effort	Academic self-efficacy Comprehension
Psychology	.418	.174	.380	.001	.008	-.024	-.002	.005	.041	.004	.065	-.049
Law	.305	.204	.263	.048	.011	-.018	-.007	.001	.006	-.072	.019	.009
Criminology	-.009	.094	.284	-.012	-.021	-.060	.006	.006	.015	.205	.010	-.116
Linguistics and literature	.121	.117	.339	.002	.013	-.043	.017	.005	.060	-.017	.033	-.035
History	.201	.249	.098	-.002	.018	-.023	-.040	-.011	.059	.123	-.020	-.054
Applied linguistics	.314	.043	.260	.005	.016	-.045	.040	-.009	.045	.044	-.091	-.007
Veterinary medicine	.167	.048	.090	.019	-.007	-.010	.009	.023	-.010	-.131	.118	-.072
Pharmaceutical sciences	.396	.340	.206	-.003	.023	-.008	-.014	.007	.005	-.046	.076	-.087
Total	.238	.175	.269	.002	.007	-.024	.004	.002	-.006	.032	.052	-.046

Table 7 Significant variables in the prediction of Passing. ($p < .05$)

Program	Hours of math in SE	Mathematics	Reading comprehension	Vocabulary	Conscientiousness	Cognitive test anxiety	Metacognition (knowledge)	Metacognition (Regulation)	Controlled motivation	Autonomous motivation	Academic self-efficacy Effort	Academic self-efficacy Comprehension
Psychology	x	x	x			x			x		x	x
Law	x	x		x						x		
Criminology						x				x		x
Linguistics and literature		x	x	x		x			x			
History		x								x		
Applied linguistics						x						
Veterinary medicine										x		x
Pharmaceutical sciences	x	x										x
Total	x	x	x	x	x	x			x		x	x

141

Prediction of GPA

To allow for comparison with the literature, Table 8 shows the variance explained in GPA for each program and for each cluster of variables (ß's are shown in Table 9). A regression analysis using the total sample yielded significant results for all variable clusters, which supported our second hypothesis. The explained variance was .171. Confirming hypothesis 5, program-specific analyses generated higher explained variances, varying between .101 and .287 with an average of .234. In all study programs, except Veterinary medicine, more than one cluster of variables significantly added to the prediction of academic success. In half of the study programs, the non-cognitive motivational/test anxiety cluster was significant (with ΔR^2 between .05 and .18).

Table 8 (increase in) explained variance for each cluster of variables (linear regression with GPA)

Program	R^2 background	ΔR^2 Cognitive skills	ΔR^2 Conscientiousness	ΔR^2 Motivation and test anxiety	Total R^2
Psychology	.083**	.056**	.016**	.045**	.201
Law	.122**	.088**	.011*	.019	.241
Criminology	.046**	.027	.010	.181**	.271
Linguistics and literature	.020**	.078**	.052**	.062**	.216
History	.003	.173**	.036*	.063	.272
Applied linguistics	.111**	.105**	.038*	.033	.287
Veterinary medicine	.030*	.014	.009	.048	.101
Pharmaceutical sciences	.118**	.058**	.025*	.080**	.280
Average	.067	.075	.025	.066	.234
Total	.049**	.075**	.016**	.031**	.171

Note. **p < .01 *p < .05

Table 9 Standardized ß Coefficients of Linear Regressions

Program	Hours of math in SE	Mathematics	Reading comprehension	Vocabulary	Conscientiousness	Test anxiety	Metacognition (knowledge)	Metacognition (Regulation)	Controlled motivation	Autonomous motivation	Academic self-efficacy Effort	Academic self-efficacy Comprehension
Psychology	.245	.191	.152	.036	.064	-.145	.001	.014	.116	.037	.135	-.166
Law	.256	.212	.103	.131	.165	-.080	-.043	.015	.030	-.118	.044	-.069
Criminology	.203	.136	.093	-.005	-.121	-.123	.148	-.025	.035	.463	-.019	-.278
Linguistics and literature	.124	.232	.160	.094	.165	-.210	.054	-.029	.128	-.001	.118	-.219
History	.046	.271	-.027	-.245	.197	.021	.151	-.281	.189	.103	.051	-.037
Applied linguistics	.242	.175	.212	.136	.135	-.171	.122	-.084	.054	.093	-.017	-.121
Veterinary medicine	.142	.071	.067	.055	-.042	.032	.100	.087	-.028	-.160	.194	-.124
Pharmaceutical sciences	.261	.145	.074	.115	.146	-.221	-.249	.117	.075	-.087	.202	-.168
Total	.170	.257	.119	.121	.065	-.116	.025	-.018	.093	.029	.124	-.160

Classification Success

As the aim of the SIMON project is to identify prospective students with very low chances of success, classification success was examined. A new set of regressions were run. First, two logistic regressions were run for each program: one which included the significant variables (shown in Table 7) as identified for prediction of passing in the total sample, and a second regression which included the significant variables for the program specific prediction. Next, the predicted membership (pass/fail) from both regressions was compared to the actual pass/fail in the program, which resulted in a total sample and a program specific classification success rates. Table 10 shows these rates for each program. Classification success was higher for the program-specific prediction ($M = 79.1$) as opposed to the total sample prediction ($M = 76.7$). Thus, using a program-specific prediction, 79.1% of the students that are predicted as failing the program will effectively fail the program, which is 2.4% higher than when using a prediction based on parameter estimates across study programs. This again supports our hypothesis 5.

Table 10 Successful Classification of Failing Students Based on Total Sample Parameter Estimates Versus Based on Program Specific Parameter Estimates

	Successful classification of failing students for prediction across study programs	Successful classification of failing students for program specific prediction
Psychology	77.4	79.5
Law	85.9	81.6
Criminology	79.4	83.5
Linguistics and literature	65.6	75.8
History	86.7	84.7
Applied linguistics	70.4	79.8
Veterinary medicine	66.7	71.2
Pharmaceutical sciences	81.1	76.5
Total	76.7	79.1

Successful Identification of At-risk Students

A classification success of 79.1% indicates that 20.9% of the at-risk students would still succeed in passing their first year of studying. Yet, in light of the open access policy it is the ambition of SIMON to minimize false negatives. Therefore, it is important that a classification cut-off is chosen that generates a high sensitivity. Currently a sensitivity of 95% is chosen as acceptable, which corresponds to a maximum of 5% of at-risk students that would unjustly get a warning that their studies are difficult to attain.

Using this 95% sensitivity to select the corresponding cut-off, 3.7% of the failing students were identified as at-risk based on the total sample prediction. In contrast, by using program specific predictions, 13.4% of the failing students could be identified. Table 11 shows these percentages for each program. Thus, using a program-specific prediction more students can be correctly identified as at-risk which again supports our hypothesis 5.

Table 11 Percentage of students correctly identified as risk-students based on total sample versus based on program specific prediction with a cut-off at a sensitivity of 95%

	% of failing students that is correctly identified based on total sample prediction with a cut-off at 95% sensitivity	% of failing students that is correctly identified based on program specific prediction with a cut-off at 95% sensitivity
Psychology	3%	6.2%
Law	3.5%	26.5%
Criminology	9.8%	26.5%
Linguistics and literature	2.7%	6.8%
History	4.2%	17.6%
Applied linguistics	5.1%	1.4%
Veterinary medicine	3.8%	3.1%
Pharmaceutical sciences	1.6%	18%
Total	3.7%	13.4%

Path Analysis and Structural Invariance

To test hypothesis 4, which implied that all variables would predict academic achievement and that test anxiety would be predicted by motivation, self-efficacy and conscientiousness, we ran path analyses using Lavaan (Rosseel, 2012) in R. Insignificant paths were deleted until a final model was reached. We started with a model in which all variables predicted GPA, and in which self-efficacy, motivational, metacognitive and personality variables predicted the academic emotion test anxiety. First, metacognition was excluded because of insignificance with both test anxiety and GPA, which of course paralleled findings from previous regressions. Self-efficacy: effort did not predict test anxiety, but motivational factors and conscientiousness did. Autonomous motivation predicted GPA through test anxiety, but not GPA directly. 20% of the variance in test anxiety was explained and 17% of the variance in GPA. The final model with standardized regression coefficients is shown in Figure 1. The model showed good fit, as indicated by χ^2 (6, 16.43), $p = .01$; RMSEA = .03 (CI .01 - .05), CFI = . 99 and NFI = .98 (Tabachnick & Fidell, 2007).
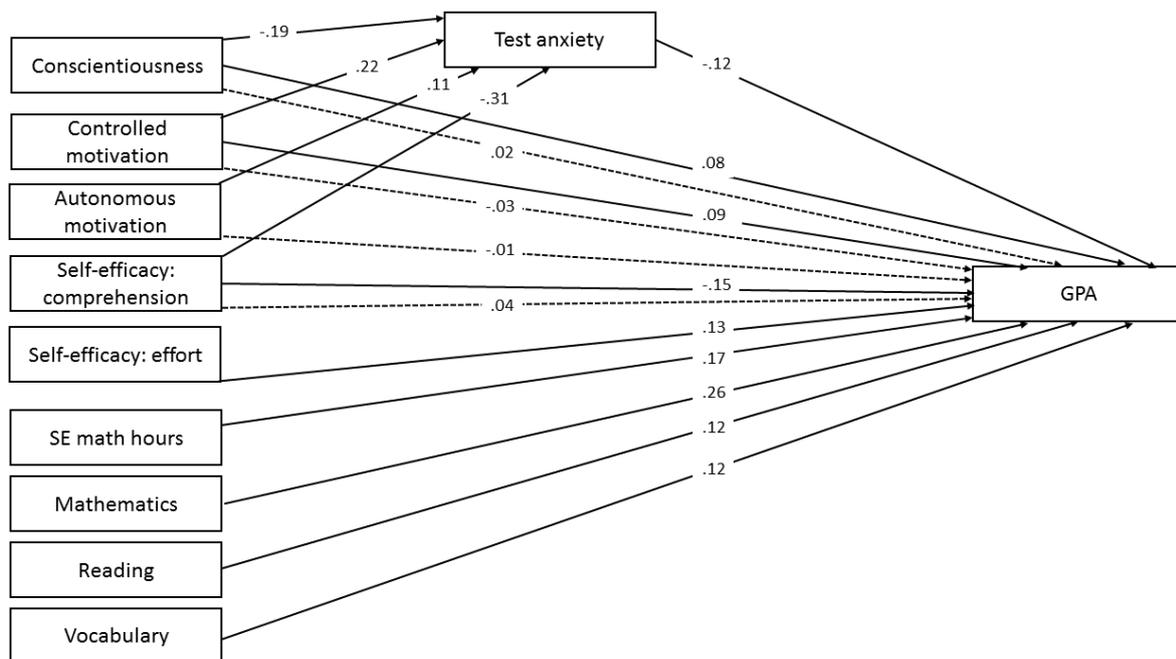

Fig. 1. Path model with standardized estimates. Only significant paths are shown. Full lines indicate direct, dotted lines indirect effects.

Next, we tested this final model for structural invariance across study programs. First, this baseline model was applied to each study program separately. Results suggested that the model did not fit all programs equally well (e.g., RMSEA criminology = .17). Finally, we conducted a multi-group analysis which involved comparing the baseline model with a second model that is constrained so that the paths are equal between groups. Since we propose that factors differentially predict academic achievement across programs, we expect the model to show structural variance. For model comparisons, we used χ^2 difference tests and looked for changes in RMSEA and CFI. The chi square difference test was significant (p < .001) and both RMSEA and CFI worsened (from .028 to .062 and from .988 to .917 respectively). This again supported our hypothesis 5 that the parameter estimates varied across study programs.

Discussion

The objective of our study was to examine the incremental predictive validity of background, cognitive, personality, metacognitive, self-efficacy and motivational factors for academic achievement in a sample that is less hindered by restriction of range and to study whether this predictive power varies across academic study programs.

As hypothesized, background and cognitive factors were predictive of academic achievement (explaining respectively six and eight percent of the variance in passing). Also, for most academic disciplines cognitive predictors and background factors as well as non-cognitive predictors (conscientiousness and self-efficacy/motivation/test anxiety) significantly explained a part of the variance in academic achievement. In three programs (Applied linguistics, Criminology and Veterinary medicine) only non-cognitive factors were predictive of passing the first year.

Although results in the first two mentioned programs may be less stable than those for groups with larger samples, results show that the inclusion of non-cognitive factors allows for better prediction of academic achievement in several programs. For admission decisions,

generally only cognitive variables are tested. These variables explain on average 12% of the variance in academic performance (Kuncel, 2010). In the current study, the combination of cognitive with non-cognitive variables explained on average 23% of the variance in GPA and passing, which corresponds to what Robbins et al. (2004) found in their meta-analysis. This increase in explained variance supports the inclusion of non-cognitive variables for orientation and admission decision (see also Kyllonen, 2012). Still, an important counter-indication is that non-cognitive variables, especially when measured through self-evaluation questionnaires and when testing is high stakes, are highly susceptible to socially-desirable responses. This problem is far less manifest when the test is used for study orientation and not for selection purposes, as is the case here, in Flanders. Yet, even in selective environments, non-cognitive variables could increase student success when used post-enrollment for assisting high-risk students (Allen et al., 2009).

The incremental validity of non-cognitive factors for academic performance varied across study programs. The significant variance explained by motivation, self-efficacy and test anxiety factors varied between 2.1% and 17.6%. In comparison, Credé and Kuncel (2008) found incremental variances between 4 and 12% and Robbins et al. (2004) found an increase of 4% over and above traditional predictors.

One may wonder whether an extra 2% in explained variance is meaningful. Allen et al. (2009) recommended to evaluate this in respect to the practical utility of the test scores. A contribution of 2% may be considered relevant when this can aid alleviating academic success and retention, whether this is through adequate study orientation and admission, or through remedial activities after enrolment. The same applies for the increase in classification success of 2.4% based on program specific prediction as opposed to prediction based on total sample parameter estimates. An increase in accuracy by 2.4% is considerable when one deals with prospective students on the verge of a life-altering study choice, especially when the wrong

choice implies considerable motivational and financial consequences, both for the individual as for society, in publicly funded education.

Moreover, in line with the ambition of the SIMON project and the open access policy, we chose to minimize the amount of respondents that are falsely identified as risk-student by selecting a classification cut-off that corresponds to a 95% sensitivity. In comparison with a total sample prediction and cut-off, 9.7% more failing students were correctly identified as being at-risk using program-specific predictions.

The variability in predictive power across study programs was also confirmed by testing a path model of relations between variables. We first tested relations as proposed in the control-value theory of achievement emotions (Pekrun, 2006). In support of this model, we did find that all variables (except metacognition) predicted GPA. Also, achievement emotion (in this study test anxiety) was influenced by motivation and by the comprehension dimension of self-efficacy (but not the effort dimension). In addition to the Pekrun model, we found that test anxiety was also affected by conscientiousness. Confirming our hypothesis of variability across study programs, the final model showed structural variance. This indicates that the structural relationship between variables differed depending on the study program. Future studies could focus more on this variability. Several authors have argued that student performance is multidimensional (Kuncel, Hezlett, & Ones, 2001; Oswald, Schmitt, Kim, Ramsay, & Gillespie, 2004). It would be interesting to thoroughly examine how and why the explanatory value of these dimensions varies by major.

Although their predictive power varied across study programs, most variables did, as expected, significantly contribute to the prediction of academic performance. Cognitive ability, conscientiousness and test anxiety predicted academic achievement in all programs, with the notable exception of Veterinary medicine. In this program, only autonomous motivation and self-efficacy (comprehension) significantly predicted passing, and both did so negatively.

Several studies have emphasized that non-cognitive constructs are critical for success in veterinary medicine (see e.g., Lewis & Klausner, 2003). Our study seems to support this claim. The negative relation between autonomous motivation and academic success is somewhat contrary to expectations, but not completely incomprehensible. Ilgen et al. (2003) found that one of the strongest motivators for choosing a career in veterinary medicine was having a pet. Although testifying of autonomous motivation, having a pet does not seem the most solid basis to succeed in an educational program. Especially not when this is combined with a restricted knowledge of the veterinary profession, which was also found by the authors, even in their selective study context. An alternative explanation is that autonomously motivated students neglect boring topics in favor of preferred ones which jeopardizes their exam performance (Senko & Miles, 2008).

The negative relation between self-efficacy (comprehension) and achievement was not a surprise and was also found in other programs. The two dimensions of self-efficacy predicted achievement differently with the comprehension dimension having negative effects, whereas the dimension of effort showed a positive relation with achievement. This is in line with Vancouver and Kendall (2006), who reason that high self-efficacy can lead to diminished effort which negatively affects performance. Our results show that it may be important to distinguish effort from comprehension when discussing academic self-efficacy, with the latter including a potential risk to overestimate one's personal abilities. Future studies need to look into this further.

Only one variable, metacognition, failed to contribute to the prediction of academic achievement in all of the study programs, similar to Kitsantas et al. (2008) and Sperling, Howard, Staley, and DuBois (2004). One possible explanation is that this is a measurement artefact. Metacognition was administered using a self-evaluation questionnaire, while Veenman, Van Hout-Wolters, and Afflerbach (2006) showed that scores on questionnaires

hardly correspond to actual behavioral measures of metacognition during task performance. Think-aloud-protocols would be an alternative, but these are time-consuming and difficult to include in an online assessment.

Most participants in studies on prediction of academic achievement have previously been selected for admission using admission tests, often heavily relying on intelligence tests (Sedlacek, 2011). Therefore, results need to be corrected for range restriction effects. In the current study, subjects have not been subjected to an admission process since all included academic study programs are open to any student who has a secondary education qualification. Moreover, the fact that a majority of students also fails the enrolled program illustrates that incoming students have more heterogeneous cognitive abilities than most U.S. samples. Yet, there is definitely a self-selection process. Of all secondary education graduates, about 63% attend tertiary education and 60% of these students enroll in an academic study program (Van Daal et al., 2013). The current results hence speak only for students who enter higher education in a completely open system, but not for the entire population per se.

The current study also has some limitations. First, although research has shown that especially conscientiousness is incrementally predictive of academic performance (Conard, 2005; de Koning et al., 2012; Farsides & Woodfield, 2003; Noftle & Robins, 2007; Poropat, 2009; Trapmann et al., 2007; Trautwein, Ludtke, Roberts, Schnyder, & Niggli, 2009) other (Big Five) personality traits were not included in the study. Vedel et al. (2015) already showed how the predictive validity of personality traits differs across study programs. Future research should examine the differential and incremental validity of other personality traits. Second, apart from personality, inclusion of other variables might augment prediction accuracy. Although 23% of variance in academic achievement was accounted for, a lot remains unexplained which calls for inclusion of other constructs. To name but a few, self-control (see e.g., Tangney et al., 2004) and other motivational constructs such as the utility value of the course (Eccles & Wigfield,

2002) have been shown to predict academic achievement. It may also be worthwhile to examine academic emotions other than test anxiety, both positive and negative, such as enjoyment or boredom (Detmers et al., 2011; Pekrun, Goetz, Titz, & Perry, 2002). Including these and other factors may allow a better prediction and thus a more comprehensive model of (program-specific) academic achievement. Third, although a range of programs were included in the current study, STEM (Science, technology, engineering and mathematics) programs were not. Future studies should test whether similar cognitive and non-cognitive constructs predict academic success in STEM areas or whether it is more beneficial to rely on more program specific knowledge. Finally, only first year academic success was predicted. Although it has been documented that first year results are powerful predictors of overall academic achievement (de Koning et al., 2012) and college retention (Allen, 1999), follow-up studies should examine whether the results hold as to timely graduation and other performance indicators.

The current study has several practical implications. The fact that non-cognitive factors have incremental predictive validity for academic outcomes over and above cognitive abilities has repercussions for admission decisions. Where possible, they should be used in admission processes and especially in study orientation. During this orientation phase, it is in the interest of the prospective student to answer honestly during non-cognitive assessments, as this would generate the most suitable advice. As such, social desirability issues stemming from non-cognitive self-tests are diminished. Indeed, the current study showed that it is possible to use self-evaluation questionnaires in an online format to assess self-regulation and motivational variables and that scores on these measures increase prediction accuracy. These assessments are relatively cheap, especially compared to labor-intensive selection procedures that intend to capture these variables such as carrying out interviews and screening letters of recommendation or essays. Therefore, it may be worthwhile to examine whether they would hold in selective contexts. In any case, they seem suitable to include in self-assessment instruments for study

orientation and for prediction of academic achievement such as SIMON. The use of instruments that assess personal abilities and the fit with educational programs can be an important leverage to increase student retention. At the very least, it enables an informed choice. In a system with open access to virtually all majors, this should encourage students to choose a program that maximizes their chance of success.

The inclusion of non-cognitive variables also opens possibilities for institutions. It allows the use of test scores for the identification of students at risk of academic failure and it facilitates the design of interventions. For example, research has shown that self-regulation training interventions can increase academic performance (Credé & Kuncel, 2008). Self-efficacy and motivation interventions can also be implemented by higher education institutions (Kitsantas et al., 2008).

The differential predictive validity of specific cognitive and non-cognitive factors across study programs also has implications for research and for counselling. Shaw et al. (2012) suggested that this variability across programs might be a consequence of the nature of the course work by major, the academic "culture" of the different majors (e.g., male-dominated or highly competitive) and of differences in grading practices. In the current study, the nature of the course work and the passing criterion were fairly uniform across programs, but more research is definitely needed on the reasons for differential predictive validity.

In anticipation of future research, investigators should be aware of the limitations of the use of subjects from specific fields of studies in predicting academic outcomes. Study samples are often constituted by psychology students as these are a convenient sample to many scholars in this research area (Busato et al., 2000; Cassady & Johnson, 2002; Chamorro-Premuzic et al., 2006; de Koning et al., 2012; Harackiewicz, Barron, Tauer, & Elliot, 2002; Komarraju & Nadler, 2013; Ridgell & Lounsbury, 2004; Ziegler, Knogler, & Buehner, 2009, to mention but a few). Many studies do not even mention the specific major of their participants (e.g., Farsides

& Woodfield, 2003), or do not take this information into account when interpreting study results (Chapell et al., 2005). Therefore, researchers should replicate findings across student populations and they should at least mention the specific study program their subjects were taken from and limit their conclusions to this program.

As for career counseling, test results should always be interpreted in light of a specific study program, and not only with regards to the study level. This study shows that although a uniform test battery is used, it is possible and valuable to make context specific predictions.

References

Ackerman, P., & Heggestad, E. (1997). Intelligence, personality, and interests: evidence for overlapping traits. *Psychological Bulletin, 121*(2), 219.

Allen, D. (1999). Desire to finish college: an empirical link between motivation and persistence. *Research in Higher Education, 40*, 461-485.

Allen, J., Robbins, S., & Sawyer, R. (2009). Can measuring psychosocial factors promote college success? *Applied Measurement in Education, 23*(1), 1-22.

Bailey, T., & Phillips, L. (2015). The influence of motivation and adaptation on students' subjective well-being, meaning in life and academic performance. *Higher Education Research & Development*, 1-16.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Boekaerts, M. (1995). The interface between intelligence and personality as determinants of classroom learning. In D. H. Saklofske & M. Zeidner (Eds.), *International handbook of personality and intelligence* (pp. 161-183). New York: Springer.

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

Busato, V., Prins, F., Elshout, 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*, 1057-1068.

Byron, K., & Khazanchi, S. (2011). A meta-analytic investigation of the relationship of state and trait anxiety to performance on figural and verbal creative tasks. *Personality and Social Psychology Bulletin, 37*(2), 269-283. doi:10.1177/0146167210392788

Cassady, J. (2004). The influence of cognitive test anxiety across the learning–testing cycle. *Learning and Instruction, 14*(6), 569-592. doi:10.1016/j.learninstruc.2004.09.002

Cassady, J., & Finch, W. (2015). Using factor mixture modeling to identify dimensions of cognitive test anxiety. *Learning and Individual Differences, 41*, 14-20. doi:10.1016/j.lindif.2015.06.002

Cassady, J., & Johnson, R. (2002). Cognitive Test Anxiety and Academic Performance. *Contemporary Educational Psychology, 27*(2), 270-295. doi:10.1006/ceps.2001.1094

Chamorro-Premuzic, T., Furnham, A., & Ackerman, P. L. (2006). Incremental validity of the typical intellectual engagement scale as predictor of different academic performance measures. *Journal of personality assessment, 87*(3), 261-268. doi:10.1207/s15327752jpa8703_07

Chapell, M. S., Blanding, Z. B., Silverstein, M. E., Takahashi, M., Newman, B., Gubi, A., & McCann, N. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*(2), 268. doi:10.1037/0022-0663.97.2.268

Chemers, M., Hu, L., & Garcia, B. (2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology, 93*(1), 55-64. doi: 10.1037//0022-0663.93.1.55

Choi, N. (2005). Self-efficacy and self-concept as predictors of college students' academic performance. *Psychology in the Schools, 42*(2), 197-205. doi:10.1002/pits.20048

Conard, M. (2005). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality, 40*(3), 339-346. doi:10.1016/j.jrp.2004.10.003

Credé, M., & Kuncel, N. (2008). Study habits, skills, and attitudes. The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science, 3*(6), 425-453. doi:10.1111/j.1745-6924.2008.00089.x

Dawis, R.V. (2005). The Minnesota theory of work adjustment. In S.D. Brown & R.W. Lent (Eds.) Career development: Putting theory and research into practice. Hoboken, NJ: Wiley.

De Backer, L., Van Keer, H., & Valcke, M. (2012). Exploring the potential impact of reciprocal peer tutoring on higher education students' metacognitive knowledge and regulation. *Instructional Science, 40*(3), 559-588. doi:10.1007/s11251-011-9190-5

De Fruyt, F., & Rolland, J. P. (2010). *PfPI. Beschrijving persoonlijkheid op het werk. Handleiding.* Amsterdam/Antwerpen: Pearson.

de Koning, B., Loyens, S., Rikers, E., Smeets, G., & van der Molen, H. (2012). Generation Psy: Student characteristics and academic achievement in a three-year problem-based learning bachelor program. *Learning and Individual Differences, 22*(3), 313-323. doi:10.1016/j.lindif.2012.01.003

De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality, 10*(5), 303-336. doi:10.1002/(SICI)1099-0984(199612)10:5%3C303::AID-PER262%3E3.0.CO;2-2

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry, 11*(4), 227-268. doi:10.1207/S15327965PLI1104_01

Detmers, S., Trautwein, U., Lüdtke, O., Goetz, T., Frenzel, A. C., & Pekrun, R. (2011). Students' emotions during homework in mathematics: Testing a theoretical model of antecedents and achievement outcomes. *Contemporary Educational Psychology, 36*(1), 25-35. doi:10.1016/j.cedpsych.2010.10.001

Dollinger, S. J., Matyja, A. M., & Huber, J. L. (2008). Which factors best account for academic success: Those which college students can control or those they cannot? *Journal of Research in Personality, 42*(4), 872-885. doi:10.1016/j.jrp.2007.11.007

Dweck, C. S. (1986). Motivational processes affecting learning. *American psychologist, 41*(10), 1040. doi:10.1037/0003-066X.41.10.1040

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

Elias, S., & Loomis, R. (2002). Utilizing Need for Cognition and Perceived Self-Efficacy to Predict Academic Performance1. *Journal of Applied Social Psychology, 32*(8), 1687-1702. doi:10.1111/j.1559-1816.2002.tb02770.x

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion, 7*(2), 336. doi:10.1037/1528-3542.7.2.336

Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: the roles of personality, intelligence, and application. *Personality and Individual Differences, 34*(7), 1225-1243. doi:10.1016/S0191-8869(02)00111-3

Ferla, J. (2008). *The effect of student cognitions about learning on self-regulated learning: a study with freshman in higher education.* (Phd), Ghent University.

Fonteyne, L., De Fruyt, F., Dewulf, N., Duyck, W., Erauw, K., Goeminne, K., . . . Rosseel, Y. (2015). Basic mathematics test predicts statistics achievement and overall first year academic success. *European Journal of Psychology of Education, 30*(1), 95-118. doi:10.1007/s10212-014-0230-9

Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2016). Exploring vocational and academic fields of study: Development and validation of the Flemish SIMON Interest Inventory (SIMON-I). *International Journal for Educational and Vocational Guidance* (In Press). Advance online publication. doi:10.1007/s10775-016-9327-9

Furnham, A., & Chamorro-Premuzic, T. (2004). Personality and intelligence as predictors of statistics examination grades. *Personality and Individual Differences, 37*(5), 943-955. doi:10.1016/j.paid.2003.10.016

Galyon, C., Blondin, C., Yaw, J., Nalls, M., & Williams, R. (2012). The relationship of academic self-efficacy to class participation and exam performance. *Social Psychology of Education : An International Journal, 15*(2), 233-249. doi:10.1007/s11218-011-9175-x

González, A., & Paoloni, P.-V. (2015). Perceived autonomy-support, expectancy, value, metacognitive strategies and performance in chemistry: a structural equation model in undergraduates. *Chemistry Education Research and Practice, 16*(3), 640-653. doi:10.1039/C5RP00058K

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. doi:10.1037//0022-0663.94.3.562

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research, 58*(1), 47-77. doi:10.3102/00346543058001047

Hill, K. T., & Wigfield, A. (1984). Test anxiety: A major educational problem and what can be done about it. *The Elementary School Journal, 85*(1), 105-126. doi:10.1086/461395

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Lutz, FL: Psychological Assessment Resources Inc.

Hong, E., & Karstensson, L. (2002). Antecedents of state test anxiety. *Contemporary Educational Psychology, 27*(2), 348-367. doi:10.1006/ceps.2001.1095

Ilgen, D. R., Lloyd, J. W., Morgeson, F. P., Johnson, M. D., Meyer, C. J., & Marrinan, M. (2003). Personal characteristics, knowledge of the veterinary profession, and influences on career choice among students in the veterinary school applicant pool. *Journal of the American Veterinary Medical Association, 223*(11), 1587-1594. doi:10.2460/javma.2003.223.1587

Kitsantas, A., Winsler, A., & Huie, F. (2008). Self-regulation and ability predictors of academic success during college: a predictive validity study. *Journal of Advanced Academics, 20*(1), 42-68. doi:10.4219/jaa-2008-867

Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences, 25*(0), 67-72. doi:10.1016/j.lindif.2013.01.005

Komarraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. *Learning and Individual Differences, 24*, 103-109. doi:10.1016/j.lindif.2012.12.007

Kuncel, N., & Hezlett, S. (2010). Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current Directions in Psychological Science, 19*(6), 339-345. doi:10.1177/0963721410389459

Kuncel, N., Hezlett, S., & Ones, D. (2004). Academic performance, career potential, creativity, and job performance: can one construct predict them all? *Journal of Personality and Social Psychology, 86*(1), 148-161. doi:10.1037/0022-3514.86.1.148

Kuncel, N., Hezlett, S. A., & Ones, D. S. (2001). A Comprehensive Meta-Analysis of the Predictive Validity of the Graduate Record Examinations: Implications for Graduate Student Selection and Performance. *Psychological Bulletin, 127*(1), 162-181. doi:10.1037//q033-2909.127.1.162

Kusurkar, R., Ten Cate, T., Vos, C., Westers, P., & Croiset, G. (2013). How motivation affects academic performance: a structural equation modelling analysis. *Advances in Health Sciences Education, 18*(1), 57-69. doi:10.1007/s10459-012-9354-3

Kyllonen, P.C. (2012). The Importance of Higher Education and the Role of Noncognitive Attributes in College Success. *Pensamiento Educativo. Revista de Investigación Educacional Latinoamericana*, 49(2), 84-100.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: a quick and valid lexical test for advanced learners of English. *Behavior research methods, 44*, 325-343. doi:10.3758/s13428-011-0146-0

Lent, R., Brown, S., & Larkin, K. (1986). Self-efficacy in the prediction of academic performance and perceived career options. *Journal of Counseling Psychology, 33*(3), 265. doi:10.1037/0022-0167.33.3.265

Lewis, R. E., & Klausner, J. S. (2003). Nontechnical competencies underlying career success as a veterinarian. *Journal of the American Veterinary Medical Association, 222*(12), 1690-1696. doi:10.2460/javma.2003.222.1690

Liebert, R. M., & Morris, L. W. (1967). Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychological reports, 20*(3), 975-978. doi:10.2466/pr0.1967.20.3.975

Lipnevich, A. A., & Roberts, R. D. (2012). Noncognitive skills in education: Emerging research and applications in a variety of international contexts. *Learning and Individual Differences, 22*(2), 173-177. doi:10.1016/j.lindif.2011.11.016

McGrath, C., Henham, M., Corbett, A., Durazzi, N., Frearson, M., Janta, B., . . . Schweppenstedde, D. (2014). *Higher Education Entrance Qualifications and Exams in Europe: A Comparison.* Retrieved from http://www.rand.org/pubs/research_reports/RR574

Messick, S. (1979). *Potential uses of noncognitive measurement in education*. Journal of Educational Psychology, 71(3), 281-292. doi: 10.1037/0022-0663.71.3.281

Ministerie van Onderwijs en Vorming (2009). Studiesucces Generatiestudenten in 2007-2008. Retrieved from

http://www.ond.vlaanderen.be/hogeronderwijs/werken/studentadmin/studentengegevens/Studi esucces%20generatiestudenten%20in%202007.pdf

Ministerie van Onderwijs en Vorming (2014). Hoger Onderwijs in cijfers -Addendum. Retrieved from http://ond.vlaanderen.be/hogeronderwijs/werken/studentadmin/studentengegevens/HOIC_201 4.pdf

Morgan, R. (1990). Predictive validity within categorizations of college students: 1978, 1981, and 1985. *ETS Research Report Series, 1990*(1), i-17.

Morris, L. W., Davis, M. A., & Hutchings, C. H. (1981). Cognitive and emotional components of anxiety: Literature review and a revised worry–emotionality scale. *Journal of Educational Psychology, 73*(4), 541. doi:10.1037//0022-0663.73.4.541

Müller, W. (2014). Educational inequality and social justice: Challenges for career guidance. *International Journal for Educational and Vocational Guidance, 14*(1), 21-33. doi:10.1007/s10775-014-9264-4

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. doi:10.1037/0022-0167.38.1.30

Murtaugh, P. A., Burns, L. D., & Schuster, J. (1999). Predicting the retention of university students. *Research in Higher Education, 40*(3), 355-371. doi:10.1023/A:1018755201899

Noftle, E., & Robins, R. (2007). Personality predictors of academic outcomes: big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*(1), 116. doi:10.1037/0022-3514.93.1.116

OECD. (2003). Career guidance: new ways forward. In OECD (Ed.), *Education policy analysis*.

OECD. (2008). Education at a Glance 2008: OECD indicators. OECD Publishing. Retrieved from http://www.oecd.org/edu/skills-beyond-school/41284038.pdf

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

Oswald, F., Schmitt, N., Kim, B., Ramsay, L., & Gillespie, M. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89*(2), 187-207.

Owen, S., & Froman, R. (1988). *Development of a college academic self-efficacy scale.* Paper presented at the National Council on Measurement in Education, New Orleans, LA.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4). doi:10.3102/00346543066004543

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341. doi:10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91-105. doi:10.1207/S15326985EP3702_4

Petrides, K. V., Chamorro-Premuzic, T., Frederickson, N., & Furnham, A. (2005). Explaining individual differences in scholastic behaviour and achievement. *British Journal of Educational Psychology, 75*(2), 239-255. doi:10.1348/000709904X24735

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

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

Ridgell, S., & Lounsbury, J. (2004). Predicting academic success: General intelligence, "big five" personality traits, and work drive. *College Student Journal, 38*(4).

Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288. doi:10.1037/0033-2909.130.2.261

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

Roth, B., Becker, N., Romeyke, S., Schafer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*, 118-137. doi:10.1016/j.intell.2015.09.002

Ryan R., C. J. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57*, 749-761.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460-475. doi:10.1006/ceps.1994.1033

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist, 32*(4), 195-208. doi:10.1207/s15326985ep3204_1

Sedlacek, W. (2010). Noncognitive measures for higher education admissions. *International encyclopedia of education, 3*, 845-849.

Sedlacek, W. (2011). Using noncognitive variables in assessing readiness for higher education. *Readings on Equal Education, 25*, 187-205.

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research, 4*(1), 27-41. doi:10.1080/08917779108248762

Senko, C., & Miles, K. M. (2008). Pursuing their own learning agenda: How mastery-oriented students jeopardize their class performance. *Contemporary Educational Psychology, 33*(4), 561-583. doi:10.1016/j.cedpsych.2007.12.001

Shaw, E. J., Kobrin, J. L., Patterson, B. F., & Mattern, K. D. (2012). *The validity of the SAT for predicting cumulative grade point average by college major*. Retrieved from http://research.collegeboard.org/publications/validity-sat-predicting-cumulative-grade-point-average-college-major

Spada, M. M., & Moneta, G. B. (2014). Metacognitive and motivational predictors of surface approach to studying and academic examination performance. *Educational Psychology, 34*(4), 512-523. doi:10.1080/01443410.2013.814196

Sperling, R. A., Howard, B. C., Staley, R., & DuBois, N. (2004). Metacognition and self-regulated learning constructs. *Educational Research and Evaluation, 10*(2), 117-139. doi:10.1076/edre.10.2.117.27905

Stark, J. S., & Lowther, M. A. (1988). Responsive professional education: Balancing outcomes and opportunities. *Studies in Higher Education, 13*(2), 223.

Sternberg, R., Grigorenko, E., & Bundy, D. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly, 47*(1), 1-41. doi:10.1353/mpq.2001.0005

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. New York: Routledge.

Swanson, J. & Schneider, M. (2013). Minnesota theory of work adjustment. In S.D. Brown & R.W. Lent (Eds.) Career Development and Counseling: Putting Theory and Research to Work, 2nd Edition. Hoboken, NJ: Wiley.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson Education.

Tangney, J., Baumeister, R., & Boone, A. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271-322. doi:10.1111/j.0022-3506.2004.00263.x

Taylor, G., Jungert, T., Mageau, G., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: the unique role of intrinsic motivation. *Contemporary Educational Psychology, 39*(4), 342-358. doi:10.1016/j.cedpsych.2014.08.002

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

Trautwein, U., Ludtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: Conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology, 97*(6), 1115-1128. doi:10.1037/a0017048

Van Daal, T., Coertjens, L., Delvaux, E., Donche, V., & Van Petegem, P. (2013). *Klaar voor hoger onderwijs of arbeidsmarkt? Longitudinaal onderzoek bij laatstejaarsleerlingen secundair onderwijs*. Antwerpen - Apeldoorn: Garant.

Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology, 91*(5), 1146. doi:10.1037/0021-9010.91.5.1146

Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology, 87*(3), 506. doi:10.1037//0021-9010.87.3.506

Vanderstoep, S. W., Pintrich, P. R., & Fagerlin, A. (1996). Disciplinary differences in self-regulated learning in college students. *Contemporary Educational Psychology, 21*(4), 345-362.

Vansteenkiste, M., Sierens, E., Soenens, B., Luyckx, K., & Lens, W. (2009). Motivational profiles from a self-determination perspective: the quality of motivation matters. *Journal of Educational Psychology, 101*(3), 671-688. doi:10.1037/a0015083

Vansteenkiste, M., Zhou, M., Lens, W., & Soenens, B. (2005). Experiences of autonomy and control among Chinese learners: Vitalizing or immobilizing? *Journal of Educational Psychology, 97*(3), 468-483. doi:10.1037/0022-0663.97.3.468

Vedel, A., Thomsen, D. K., & Larsen, L. (2015). Personality, academic majors and performance: Revealing complex patterns. *Personality and Individual Differences, 85*, 69-76. doi:10.1016/j.paid.2015.04.030

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning, 1*(1), 3-14. doi:10.1007/s11409-006-6893-0

Vuong, M., Brown-Welty, S., & Tracz, S. (2010). The effects of self-efficacy on academic success of first-generation college sophomore students. *Journal of College Student Development, 51*(1), 50-64.

Wiberg, M., & Sundström, A. (2009). A comparison of two approaches to correction of restriction of range in correlation analysis. *Practical Assessment, Research & Evaluation, 14*(5), 2.

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

Ziegler, M., Knogler, M., & Buehner, M. (2009). Conscientiousness, achievement striving, and intelligence as performance predictors in a sample of German psychology students: Always a linear relationship? *Learning and Individual Differences, 19*(2), 288-292. doi:10.1016/j.lindif.2009.02.001

Chapter 6: Career goal engagement following negative feedback: influence of expectancy-value and perceived feedback accuracy[5]

Abstract

What happens when students receive feedback that the study major they envisioned is difficult to attain? Either they continue goal engagement but adapt behavior (assimilation), or they abandon the goal (accommodation). Thus, for the former students negative feedback is an encouragement to double their efforts whilst for the latter it prompts disengagement and exploration of other opportunities. Little is known on the effects of negative attainability feedback for these goal management strategies when choosing a study major. What processes are in play when shifting from engagement to disengagement from a career goal? These issues are addressed in the current study by drawing on expectancy-value theory, goal-setting models and the dual-process framework. More specifically, it was hypothesized that receiving negative attainability feedback may lead to both goal engagement (assimilation) and goal disengagement (accommodation) and that this relation is mediated by self-efficacy, motivational beliefs and by the perceived accuracy of feedback.

Results showed that negative attainability feedback led to goal disengagement and, to a lesser extent, to continued engagement. Perceived accuracy of feedback was an important mediator for goal management, as was motivation. Contrary to expectations, self-efficacy did not predict either goal management strategy.

_____

5 Fonteyne, L., Eelbode, A., Lanszweert, I., Roels, E., Schelfhout, S., Duyck, W., & De Fruyt, F. (2017). Career goal engagement following negative feedback: influence of expectancy-value and perceived feedback accuracy. Manuscript submitted for publication.

Career goal engagement

Choosing and pursuing a career goal is often experienced as a daunting task. During the career choice process, adolescents need to take into account both their personal abilities, interests and values and weigh them against the demands of educational or job choices. They may need to compromise on a career goal because of contextual (e.g., distance, availability of study program…) or personal (e.g., intellectual, motivational…) constraints. This idea of compromise between aspirations and reality is well-embedded in theories of occupational choice such as the Theory of Circumscription and Compromise (Gottfredson, 1981) and the Career Construction Theory (Savickas, 2006). It is also a central premise of more general goal-setting models (e.g., Bandura's social cognitive theory (1991) and Carver and Scheier's control model (1990)), which state that behavior and goal-setting are guided by a feedback loop in which there is a continual evaluation of the attainability of goals. When there is a discrepancy between the desired and the actual state, the behavior and/or the goal is adapted to align them.

This management of behavior and goals is addressed in theories of developmental regulation such as the dual-process framework (Brandtstädter & Rothermund, 2002) and the Motivational theory of life-span development (Heckhausen, Wrosch, & Schulz, 2010). Both theories provide a framework to understand the dynamic processes by which goals are adapted. A discrepancy between an actual and a desired state can be reduced either by trying to change the situation to align more closely with the goal (goal engagement) or by adjusting the goal to meet the situational constraints (goal disengagement). The dual-process framework refers to the first process as assimilation and the latter as accommodation (Brandtstädter, 2009). The assimilative processes are aimed at effective goal pursuit and thus goal engagement. Yet, when there is repeated failure or when a goal becomes unattainable, accommodative processes become more beneficial.

Disengagement may benefit health (Miller & Wrosch, 2007) and well-being (Wrosch, Scheier, Carver, & Schulz, 2003). Creed and Hood (2014) showed that disengagement capacity may also have beneficial effects in the career development context. More specifically, they found that students with a higher capacity to disengage from unattainable goals experienced less career distress.

However, for a (prospective) student it is extremely difficult to evaluate whether a career goal is really unattainable and when exactly it is better to give up. For some, changing a specific career plan and switching to another study major is particularly challenging. To date, little is known on how a person goes from goal engagement to disengagement (Heckhausen et al., 2010). And notwithstanding a few exceptions, this self-regulatory challenge has received scant attention in the career context (Praskova, Creed, & Hood, 2013).

The current study addresses this gap in the literature on goal (dis)engagement in the career context. Although it is generally accepted that goal management is guided by a feedback loop, the question how negative feedback about goal attainability influences career goal (dis)engagement remains unanswered. We hypothesize that goal (dis)engagement following feedback will be mediated by expectancy and value variables and by the perceived accuracy of the feedback.

Expectancy and value

Two important factors are assumed to influence the ease of disengagement (Brandtstädter and Rothermund, 2002). First, the subjective attainability of a goal and, second, its personal importance and centrality. These factors are in line with the expectancy-value theory of achievement motivation (see e.g., Eccles & Wigfield, 2002) in which achievement choices are a function of the expectancy of success and the value of the goal, which are influenced by ability beliefs, perceived difficulty and self-schemes. These are, in turn, shaped by previous experiences and socialization (Wigfield & Eccles, 2000). In the context of career

goal management, this implies that disengaging from a career goal would be less likely when it is perceived as important and when the expectancy of success is high.

Yet, during the career choice process, a proper estimate of the likelihood of success is problematic. Students on the verge of making career and/or educational choices lack previous experiences within the higher education context, which hinders the evaluation of attainability. As such, realistic goal-setting becomes difficult. Therefore, students actively seek feedback on whether their goal is suitable and which plans and actions are appropriate (Creed, Wamelink, & Hu, 2015; Hattie & Timperley, 2007; Kerpelman, Pittman, & Lamke, 1997). They seek advice from peers, parents, teachers, career guidance counsellors or online assessment instruments.

Still, even when feedback is available, some students ignore this feedback. They persist in engaging with a career goal in spite of feedback that the goal is unattainable. Again, self-efficacy and value seem to influence the acceptance of this kind of feedback. For example, just as self-efficacy influences goal management, it has been shown to affect the reactions to negative feedback (Ilgen & Davis, 2000).

Reactions to feedback

There is ample evidence that (especially negative) feedback is difficult to accept. Feedback that is inconsistent with expectancies is often discarded. For example, Sinclair and Cleland (2007) found that low achievers were less likely to collect feedback than high achievers. Often, those most in need of feedback seem to engage the least with the feedback they receive (Harrison et al., 2013).

This is in line with a control theoretical perspective (Carver & Scheier, 1990), wherein negative feedback provokes feelings of resistance, which in turn leads people to discard the feedback. In the context of employee selection procedures for example, Schmitt, Oswald, Kim, Gillespie, and Ramsay (2004) found that poorly performing examinees evaluated the test as

invalid and irrelevant to the job. The authors explained this finding using the self-serving bias. In the case of negative feedback, the importance or the credibility of the feedback is refuted or the feedback is ignored altogether in order to protect self-worth. Thus, from a control theoretical perspective, negative feedback encourages continued goal engagement, either by discarding the feedback or by signaling that more effort is needed.

According to social cognitive theory (Bandura, 1991) on the other hand, negative feedback decreases people's confidence and thus their success expectations which leads them to disengage from the goal. And indeed, research has shown that people do lower their goals after receiving negative feedback (see e.g., Ilies & Judge, 2005; Krenn, Wurth, & Hergovich, 2013), or abandon their goal altogether (Kluger & DeNisi, 1996).

Thus, different goal-setting models predict opposing reactions to negative feedback and research findings are indeed not straightforward. Individual and situational differences seem to influence feedback reactions (Eva et al., 2012).

In the context of career guidance, it is highly relevant to study the processes involved in these reactions. Many students will come to realize that their educational goal is unattainable (Boudrenghien, Frenay, & Bourgeois, 2012). They are better off disengaging from their goal early on or preferably even before starting their study trajectory, rather than persisting unsuccessfully. Several sources provide (prospective) students with attainability feedback but this feedback is often discarded. As this study seeks to shed light on how negative attainability feedback influences career goal disengagement, its results might support the development of personalized feedback strategies that promote adequate career goal management.

Current study

The current study differs in important ways from previous research. First, although the benefits of goal disengagement have been demonstrated, the processes and factors that influence disengagement are still unclear (see Heckhausen et al., 2010). Especially studies in

ecologically valid settings are rare (Tomasik & Silbereisen, 2012). Moreover, there is limited research about disengagement from career goals during the study choice process (Creed & Blume, 2013). Also, we know little about the role that feedback plays in career development (Creed et al., 2015). Many studies on feedback reactions have focused on the impact of characteristics of feedback delivery, whereas it is equally important to consider feedback from the perspective of the receiver (Eva et al., 2012) and how it affects subsequent action and behavior. Finally, most of the feedback literature is focused on feedback that fosters performance improvement. Consequently, little is known on how to give feedback to promote disengagement from goals.

In the present study, we examine the effect of negative attainability feedback on career goal management. Can negative attainability feedback encourage career goal disengagement at the start of the university trajectory? How do students react to negative attainability feedback (as opposed to positive attainability feedback): by doubling their effort (assimilation and engagement as proposed by control theory) or by exploring other options (accommodation and disengagement as suggested in social cognitive theories)? And to what extent are these management strategies mediated by self-efficacy, motivation and the perceived accuracy of feedback?

It is hypothesized that negative feedback will lead to goal engagement in some students (control theory) and to disengagement in others (social cognitive theory). We hypothesize that whether a student continues to engage with or disengages from a career goal following negative feedback will be mediated by expectancy, value and perceived accuracy of feedback.

More specifically, it is expected that receiving negative feedback will influence self-efficacy and motivation negatively. In turn, self-efficacy and motivation will positively relate to assimilation and negatively to accommodation. Following control theory, we hypothesize that receiving negative feedback will decrease the perceived accuracy of the feedback which in

turn increases goal engagement (assimilation) and disengagement (accommodation). See Figure 1 for a graphical representation of hypothesized relations.
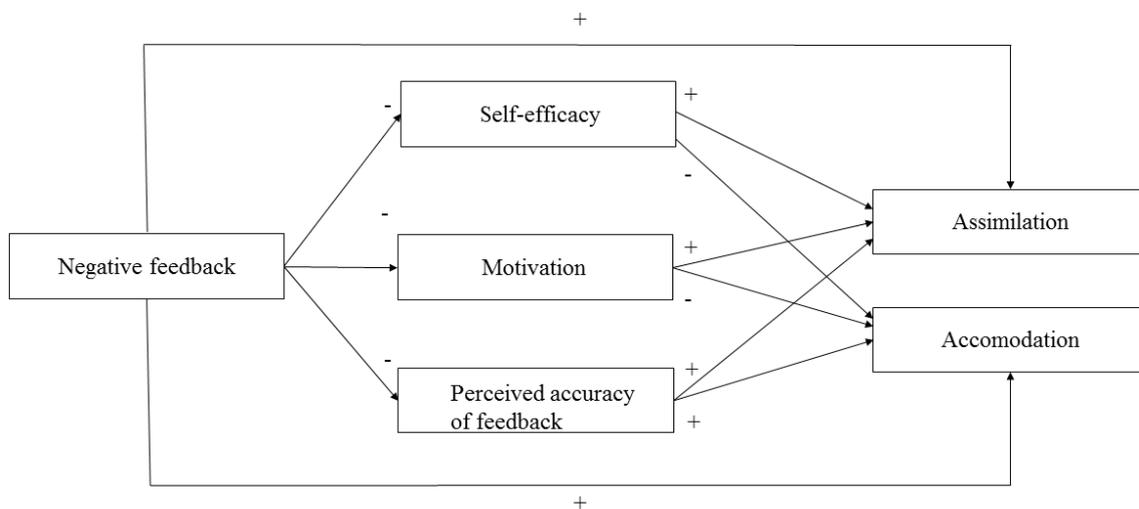

Fig. 1. Model of hypothesized relations.

At a more descriptive level, we are interested to evaluate to what extent students who received negative attainability feedback are activated by their feedback report (by putting in more effort for their studies, by participating in guidance activities or by considering to change majors).

Method

Participants and general set-up

The current study was set in Flanders, a region in Belgium where the educational system is characterized by virtually unlimited open access. Students with any secondary education qualification can enroll in almost any study major (with few exceptions such as medicine and dentistry), without specific requirements such as passing a selection exam or obtaining a certain GPA. At the start of the academic year, new incoming undergraduate students at a large Flemish

University filled out an online test battery called SIMON (Study skills and Interest MONitor). This test is aimed at the identification of a small group of students (usually about 10%) that lack the necessary skills to pass their first year of higher education (see, Fonteyne, Duyck, & De Fruyt, 2017; Fonteyne et al., 2016). Thus, in contrast with high-stake admission tests, SIMON's discriminatory power lies at the lower end of the ability range which complies with the open access policy: only students who almost certainly lack the very basic abilities to succeed get a clear warning, yet students who might be able to pass get the benefit of doubt and are not discouraged. Therefore, based on the literature on predictors of academic achievement and retention, tests of very basic skills that are necessary prerequisites to pass in higher education were selected to include in the battery (basic reasoning skills, basic mathematical knowledge, vocabulary knowledge, reading comprehension, motivation, self-efficacy, metacognition, test anxiety, self-control and grit).

Three weeks into the academic year the students received a personal feedback report. This report entailed detailed scores on each of the skills along with elaborate information on remedial courses to improve skills in order to pass first year at university. The report also provided a personalized estimate of how likely it was that the student would successfully pass that specific study program. This chance of success was previously validated using historical data of over 15,000 students in all 11 faculties of Ghent University. This empirical validation shows us that only 5% of the students who receive a low chance actually pass whereas 70% of the students with a high chance of passing succeed (for further details on the predictive validity of the test battery, see Fonteyne et al. (2017), i.e. chapter 5). 9.2% of the students were informed that chance of passing the first year was low, i.e. lower than 5%. 6.2% of students were informed that chance of passing was high, i.e. higher than 70%. This validity evidence was also included in the feedback reports (see Appendix A for the first three pages of an example report, English translations are framed). The majority of the students (84.6%) received a feedback report in

which it was clearly stated that prediction of success was difficult. These students still received their personal scores and information on remedial courses, but they were not included in the current study.

One month after receiving their personalized feedback report, students were invited to evaluate the received report. This evaluation also included measures of self-efficacy, motivation, perceived accuracy of feedback and of goal (dis)engagement. 25.7% of the students with a low chance of success ($N = 117$) and 39.4% of students with a high chance of passing ($N = 121$) responded to this invitation and were included in the current study. Thus, the final sample consisted of 238 respondents. The study was approved by the faculty ethics committee and students gave written consent for participation.

Measures

All items were measured on a 4-point scale (ranging from *totally disagree* to *totally agree*).

Self-efficacy was measured using the item 'How certain are you that you are going to pass your study program?' whereas *Value* was assessed with the item 'How important is it for you to take this study program?'.

Perceived Accuracy of Feedback was assessed using the Fairness (3 items, e.g., 'I consider this feedback fair'), the Usefulness (3 items, e.g., 'I consider this feedback helpful') and 2 items from the Acceptance subscales (e.g., reversed item 'I reject this feedback') of the Perceived Accuracy of Feedback scale (Strijbos, Narciss, & Dünnebier, 2010). Cronbach's α in the current sample was .92.

For goal engagement and disengagement, the assimilation and accommodation scale (Haratsis, Creed, & Hood, 2015) was adapted to fit the study context. Students were instructed to keep their attainability feedback in mind by queuing them with 'Because of the feedback I received…'. For each of the subscales, 10 statements followed (e.g., assimilation '…I will

double my efforts' and accommodation '…I will focus on a different study program'). Cronbach's alphas in the current sample were .88 for assimilation and .95 for accommodation.

We also asked whether students participated in study guidance activities (1 item), whether they would put more effort into their studies (1 item) and whether they had considered changing majors (1 item), as a result of the feedback report.

Results

All analyses were performed using SPSS24 and AMOS22. Means and correlations for the variables are listed in Table 1. One case was identified as a multivariate outlier through Mahalanobis distance with $p < .001$. This case was deleted, leaving 237 students for further analysis.

Table 1 Means and correlations.

	M (SD)	Assimilation	Accommodation	Self-efficacy	Motivation
Assimilation	28.26 (4.63)	1			
Accommodation	15.50 (5.42)	-.09	1		
Self-efficacy	2.42 (.66)	.12	-.12	1	
Motivation	3.37 (.60)	.25**	-.36**	.11	1
Perceived Accuracy of Feedback (PAF)	21.15 (5.10)	.32**	-.06	.22**	.12

**$p < .01$

The proposed multiple mediator model with standardized coefficients is shown in Figure 2. The model fitted the data well, as indicated by a several fit indices (Tabachnick & Fidell, 2007): χ^2 (4, 237) = 6.60, $p = .16$; CFI = .99; NFI = .98; RMSEA = .05.

As hypothesized, negative attainability feedback was significantly and negatively related to self-efficacy (unstandardized coefficient = -.25, $p < .01$) and motivation (unstandardized coefficient = -.21, $p < .01$). Yet, only 4% of the variance in expectancy and 3% of the variance in value was accounted for by negative attainability feedback. Of these variables,

only motivation had an effect on goal management strategies (unstandardized coefficients of 1.79, $p < .001$ for assimilation and -2.84, $p < .001$ for accommodation). Contrary to our expectations, self-efficacy did not significantly predict either goal management strategy (unstandardized coefficients of .23, $p = .58$ and -.64, $p = .18$ respectively). This suggests that value, but not self-efficacy is a mediator between negative attainability feedback and assimilation and accommodation.
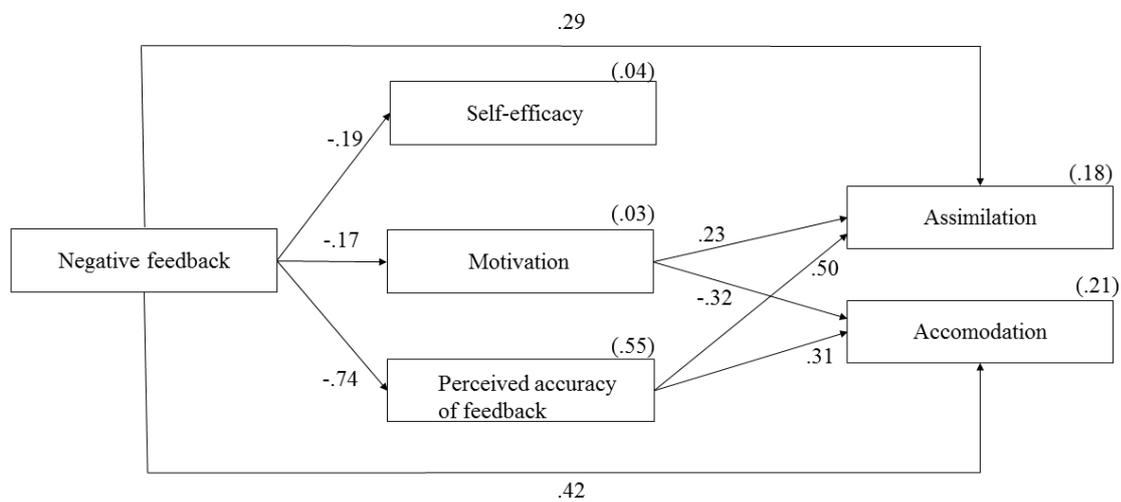

Fig. 2. Model with standardized estimates. Only significant paths are shown. Squared multiple correlations are between brackets.

Negative feedback, as expected, was negatively related to the perceived accuracy of the feedback (unstandardized coefficient = -7.59, $p < .001$). In turn, the perceived accuracy of feedback significantly predicted both assimilation and accommodation (unstandardized coefficients of .45, $p < .001$ and .33, $p < .001$ respectively). Thus, perceived accuracy of feedback potentially mediated between feedback and goal management strategies.

We further assessed the indirect effects of the proposed mediators (self-efficacy, value and perceived accuracy of feedback) following guidelines by Preacher and Hayes (2008). We

used the Amos bootstrapping procedure (1000 samples) to calculate 95% bias-corrected confidence intervals (CIs). The effect of negative attainability feedback on assimilation was fully mediated by value and perceived accuracy of feedback, as indicated by 95% CIs that did not contain zero (CI = -.91; -.07 and CI = -5.21; -2.00 respectively). Value (CI = -3.77; -.88) and perceived accuracy of feedback (CI = .21; 1.15) mediated the relationship of negative feedback with accommodation, but this relationship remained significant, indicating partial mediation. The examination of pairwise contrasts of indirect effects showed that the indirect effect through perceived accuracy of feedback was larger than the indirect effect through value (with a CI of 1.69 to 4.92 for assimilation and a CI of 1.76 to 4.53 for accommodation).

Negative attainability feedback had a significant positive direct effect on career goal management. The effect on accommodation (standardized coefficient = .42 $p < .001$) was higher than that on assimilation (standardized coefficient = .29, $p < .001$), ($F (1 , 236) = 17.76$, $p < .000$). Negative feedback had significant negative indirect effects on assimilation (standardized coefficient = -.43, $p < .001$) and accommodation (standardized coefficient = -.16, $p < .001$) through perceived accuracy of feedback.

In total, one fifth (21%) of the variance in accommodation and 18% of the variance in assimilation was accounted for by variables in the model. 55% of the variance in perceived adequacy of feedback was accounted for by the receipt of negative attainability feedback.

68.5 % of the students receiving negative feedback reported that they were activated by their feedback report. 29.1% said that they participated in study guidance activities, 15.4% had considered changing majors and 59.8% indicated they would put more effort into their studies as a result of the report.

Discussion

During the study choice process, it is very difficult for students to evaluate the suitability of a specific study program in terms of attainability. In a system with open access to study

majors, many students enroll in a program that turns out to be not attainable for them. In the current study context, 37.5% of the undergraduates drop out before the end of the first academic year (Lacante et al., 2001). These students would be better off realizing early in their academic trajectory that their chances of success are very low so they can direct their efforts towards a more suitable study program. Yet, when guidance counsellors or teachers try to communicate this message, such feedback is often discarded. In light of career counselling, it is important to know whether and when such information has an effect on career goal management. This could for example further prompt the development of personalized feedback strategies.

To investigate this, we examined whether receiving negative attainability feedback based on a validated test battery (with shown predictive validity of academic achievement) affected career goal management strategies, here operationalized as assimilation (goal engagement) and accommodation (goal disengagement). We also checked whether expectancy-value variables and the perception of feedback mediated effects on career goal management.

Following expectancy-value theory, we proposed that negative attainability feedback would lead to lower self-efficacy and lower motivation which in turn would lead to lower assimilation and higher accommodation. Negative feedback indeed had a significant negative relation with both self-efficacy and motivation. Lower motivation led to lower assimilation, and to higher accommodation. Yet, contrary to our expectations, we found no significant effect of self-efficacy on either accommodation or assimilation. Thus, whereas social cognitive theories would suggest highly self-efficacious students are more likely to persist in the presence of goal-performance discrepancies (Williams, Donovan, & Dodge, 2000), we found no such effect.

There was a direct negative effect of negative feedback on its perceived accuracy, which is in line with control theory and previous research in other contexts. The current study demonstrates that this effect is also present in the context of career goal management: when attainability feedback is negative, its relevance and credibility is affected. This effect is quite

174

strong: over half of the variance in perceived accuracy of feedback was accounted for by the receipt of negative feedback. In turn, when feedback is perceived as less accurate, this has negative effects on both accommodation and assimilation. Thus, there is a significant indirect negative effect of receiving negative feedback on goal management strategies, which is mediated by perceived accuracy of feedback.

On the other hand, the direct effect of negative feedback on both goal management strategies was positive and stronger than the indirect effects. This suggests that giving negative feedback is not futile. It does lead to action. Moreover, the effects were stronger for accommodation than for assimilation. Giving negative attainability feedback triggers enrolling students to put more effort into their studies, but even more so, it encourages them to explore other options that might be more viable for them. Either way, the effect for their own academic trajectory can be beneficial.

In sum, these results suggest that giving negative feedback does promote both goal engagement and goal disengagement. Yet, this effect is somewhat undermined by the perceived accuracy of the feedback. This makes the case for devoting special attention to how feedback is perceived, especially and specifically in the context of career goal management. A lot of the feedback research has focused on persistence or continued engagement after receiving feedback. Still, in the context of unattainable career goals it would be especially interesting to examine what feedback characteristics are important to encourage goal disengagement. For example, feedback specification may be relevant. Carroll, Shepperd, and Arkin (2009) found that students who received fully specified threatening feedback in a laboratory setting were significantly less committed to entering a fictitious program ($d = .82$) and had significant lower admission expectations ($d = 1.62$). It may also be relevant to study interactions between feedback characteristics and individual differences. For example, it is possible that highly

motivated or self-efficacious students require fully specified feedback reports to be activated, whereas highly anxious students are better off with softly formulated feedback.

Future research may also address some limitations of the current study. For self-efficacy and motivation, only one item could be used. A more thorough examination of expectancy and value beliefs may reveal other patterns. Also, the career goal management strategies were examined in students who had already enrolled in a study program. The question remains whether prospective students would react in similar ways to such negative attainability feedback. For such a group, accommodation effects may be much stronger, because they entail less cost (time, financial, logistic) relative to a choice already made. Next, students received feedback based on an online assessment of their competencies. It may also be interesting to examine to what extent the current findings apply to feedback received from relevant others such as peers, parents or teachers. And finally, because of the cross-sectional design of the current study, it is not possible to evaluate whether the attainability feedback will have long term effects. Previous research indicates that initial negative feedback leads to an increase in instead of to a withdrawal of effort (Nease, Mudgett, & Quiñones, 1999). It would be interesting to see whether and when negative feedback leads to more accommodation in the long run. For example, it is possible that receiving negative feedback at the start of higher education strengthens the effect of receiving disappointing exam results. Yet, such longer term effects require a longitudinal follow-up of students.

In any case, whether it is by doubling their efforts or by considering other study programs, the current study shows that giving negative attainability feedback does activate (68.5% of the) students early on in their academic trajectory.

References

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287. doi:10.1016/0749-5978(91)90022-L

Boudrenghien, G., Frenay, M., & Bourgeois, E. (2012). Unattainable educational goals: Disengagement, reengagement with alternative goals, and consequences for subjective well-being. *European Review of Applied Psychology-Revue Europeenne De Psychologie Appliquee, 62*(3), 147-159. doi:10.1016/j.erap.2012.04.002

Brandtstädter, J. (2009). Goal pursuit and goal adjustment: Self-regulation and intentional self-development in changing developmental contexts. *Advances in Life Course Research, 14*(1), 52-62. doi:10.1016/j.alcr.2009.03.002

Brandtstädter, J., & Rothermund, K. (2002). The life-course dynamics of goal pursuit and goal adjustment: A two-process framework. *Developmental Review, 22*(1), 117-150. doi:10.1006/drev.2001.0539

Carroll, P. J., Shepperd, J. A., & Arkin, R. M. (2009). Downward Self-Revision: Erasing Possible Selves. *Social Cognition, 27*(4), 550-578. doi:10.1521/soco.2009.27.4.550

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological review, 97*(1), 19. doi:10.1037/0033-295X.97.1.19

Creed, P. A., & Blume, K. (2013). Compromise, Well-Being, and Action Behaviors in Young Adults in Career Transition. *Journal of Career Assessment, 21*(1), 3-19. doi:10.1177/1069072712453830

Creed, P. A., & Hood, M. (2014). Disengaging From Unattainable Career Goals and Reengaging in More Achievable Ones. *Journal of career Development, 41*(1), 24-42. doi:10.1177/0894845312471195

Creed, P. A., Wamelink, T., & Hu, S. (2015). Antecedents and consequences to perceived career goal-progress discrepancies. *Journal of Vocational Behavior, 87*, 43-53. doi:10.1016/j.jvb.2014.12.001

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

Eva, K. W., Armson, H., Holmboe, E., Lockyer, J., Loney, E., Mann, K., & Sargeant, J. (2012). Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. *Advances in Health Sciences Education, 17*(1), 15-26. doi:10.1007/s10459-011-9290-7

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. Advance online publication. doi: 10.1016/j.lindif.2017.05.003

Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2016). Exploring vocational and academic fields of study: Development and validation of the Flemish SIMON Interest Inventory (SIMON-I). *International Journal for Educational and Vocational Guidance*(In Press). doi:10.1007/s10775-016-9327-9

Gottfredson, L. S. (1981). Circumscription and Compromise - A Developmental Theory of Occupational Aspirations. *Journal of Counseling Psychology, 28*(6), 545-579. doi:10.1037//0022-0167.28.6.545

Haratsis, J. M., Creed, P. A., & Hood, M. (2015). Measuring assimilative and accommodative resources in young adults: Development and initial validation of suitable scales. *Personality and Individual Differences, 81*, 61-66. doi:10.1016/j.paid.2014.12.011

Harrison, C. J., Konings, K. D., Molyneux, A., Schuwirth, L. W. T., Wass, V., & van der Vleuten, C. P. M. (2013). Web-based feedback after summative assessment: how do students engage? *Medical education, 47*(7), 734-744. doi:10.1111/medu.12209

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi:10.3102/003465430298487

Heckhausen, J., Wrosch, C., & Schulz, R. (2010). A motivational theory of life-span development. *Psychological review, 117*(1), 32. doi:10.1037/a0017668

Ilgen, D., & Davis, C. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology, 49*(3), 550-565. doi:10.1111/1464-0597.00031

Ilies, R., & Judge, T. A. (2005). Goal regulation across time: the effects of feedback and affect. *Journal of Applied Psychology, 90*(3), 453. doi:10.1037/0021-9010.90.3.453

Kerpelman, J. L., Pittman, J. F., & Lamke, L. K. (1997). Toward a Microprocess Perspective on Adolescent Identity Development An Identity Control Theory Approach. *Journal of Adolescent Research, 12*(3), 325-346. doi:10.1177/0743554897123002

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254. doi:10.1.1.461.6812

Krenn, B., Wurth, S., & Hergovich, A. (2013). The Impact of Feedback on Goal Setting and Task Performance Testing the Feedback Intervention Theory. *Swiss Journal of Psychology, 72*(2), 79-89. doi:10.1024/1421-0185/a000101

Lacante, M., De Metsenaere, M., Lens, W., Van Esbroek, R., De Jaeger, K., De Coninck, T., . . . Santy, L. (2001). *Drop-Out in het hopger onderwijs. Onderzoek naar achtergronden en motieven van drop-out in het eerste jaar hoger onderwijs.* Retrieved from Leuven-Brussel: https://cris.cumulus.vub.ac.be/portal/en/publications/dropout-in-hoger-onderwijs-onderzoek-naar-achtergronden-en-motieven-van-dropout-in-het-eerste-jaar-hoger-onderwijs(bc98955e-23a8-4625-a0f7-b910d1a73e7e).html

Miller, G. E., & Wrosch, C. (2007). You've Gotta Know When to Fold'Em Goal Disengagement and Systemic Inflammation in Adolescence. *Psychological science, 18*(9), 773-777. doi:10.1111/j.1467-9280.2007.01977.x

Nease, A. J. A., Mudgett, B. O., & Quiñones, M. A. (1999). Relationships among feedback sign, self-efficacy, and acceptance of performance feedback. *Journal of Applied Psychology, 84*(5), 806. doi:10.1037/0021-9010.84.5.806

Praskova, A., Creed, P. A., & Hood, M. (2013). Facilitating engagement in new career goals: the moderating effects of personal resources and career actions. *International Journal for Educational and Vocational Guidance, 13*(2), 115-134. doi:10.1007/s10775-013-9242-2

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods, 40*(3), 879-891. doi:10.3758/BRM.40.3.879

Savickas, M. L. (2006). *Career construction theory.* Paper presented at the AACC 15th Annual National Conference, Sydney, Australia.

Schmitt, N., Oswald, F. L., Kim, B. H., Gillespie, M. A., & Ramsay, L. J. (2004). The Impact of Justice and Self-Serving Bias Explanations of the Perceived Fairness of Different Types of Selection Tests. *International Journal of Selection and Assessment, 12*(1-2), 160-171. doi:10.1111/j.0965-075X.2004.00271.x

Sinclair, H. K., & Cleland, J. A. (2007). Undergraduate medical students: who seeks formative feedback? *Medical education, 41*(6), 580-582. doi:10.1111/j.1365-2923.2007.02768.x

Strijbos, J.-W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency? *Learning and Instruction, 20*(4), 291-303. doi:10.1016/j.learninstruc.2009.08.008

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson Education.

Tomasik, M. J., & Silbereisen, R. K. (2012). Beneficial effects of disengagement from futile struggles with occupational planning: A contextualist-motivational approach. *Developmental Psychology, 48*(6), 1785. doi:10.1037/a0026952

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

Williams, K. J., Donovan, J. J., & Dodge, T. L. (2000). Self-regulation of performance: Goal establishment and goal revision processes in athletes. *Human Performance, 13*(2), 159-180. doi:10.1207/s15327043hup1302_3

Wrosch, C., Scheier, M. F., Carver, C. S., & Schulz, R. (2003). The importance of goal disengagement in adaptive self-regulation: When giving up is beneficial. *Self and Identity, 2*(1), 1-20. doi:10.1080/15298860309021

Appendix

First three pages of a feedback report. English translations are framed.

FEEDBACKRAPPORT

NEEM DIT PERSOONLIJK RAPPORT STEEDS MEE NAAR DE STUDIE- OF TRAJECTBEGELEIDING

Please take this report with you when you visit your study counsellor

UNIVERSITEIT GENT

1 Feedback[1]

Opleiding: Pedagogische wetenschappen

Beste Voornaam,

Je vulde bij de start van het academiejaar SIMON in. SIMON wil jou helpen en

- berekent daarom jouw persoonlijke slaagkans in het eerste jaar (deel 1),
- toont jou op welke vaardigheden je goed of minder goed scoort (deel 2.A en 2.B) en
- geeft je ook aan wat je kan doen om je studies zo goed mogelijk aan te pakken (deel 2.C).

Hieronder vind je jouw persoonlijke resultaten.

Twijfel je over je opleiding? Ga langs bij je studie- of trajectbegeleider.

This report contains your personal SIMON test results .
Are you doubting your choice of program? Visit your study counsellor.

Jouw berekende SLAAGKANS is LAAG

Your personal chance of passing this program succesfully is LOW

Slechts 5% van de studenten met deze slaagkans slaagt voor het eerste jaar van deze opleiding.

Slagen betekent in dit rapport het behalen van alle opgenomen studiepunten in je eerste jaar. Deze slaagkans geldt enkel voor jouw opleiding.

Misschien is deze opleiding helemaal niets voor jou en moet je jouw studiekeuze opnieuw in vraag stellen. Maak een afspraak en praat erover met een facultaire studie- of trajectbegeleider (Meer info: http://www.ugent.be/monitoraat) of met een studieadviseur (http://www.ugent.be/Studieadvies) van de afdeling Studieadvies.

Only 5% of the students with a low chance passes the first year of this study program.
This chance is only valid for the program you are currently enrolled in.
Maybe this program does not really suit you and maybe you should reconsider your choice of study program. Make an appointment and talk about it with your faculty study counsellor (+ link to contact details).

---

[1] Een uitvoerige handleiding bij het feedbackrapport vind je op http://www.simon.ugent.be/instroom.

2   Subscores

Your personal score on each of the SIMON subtests is shown below.
You receive three kinds of information:
A. The graph shows how you score compared to the other students in your group.
B. You receive an evaluation of your skills based on research on student success: if you score low, it is advised you try to work on this skill in order to manage your studies well. If you score high, you seem to master this skill, but you are of course still welcome to join in on study guidance activities.
These 2 kinds of information can sometimes seem inconsistent. For example, you can score higher than the students in your group but still score too 'low' to have mastered this skill.
C. What next?
In this section you can find an overview of the organised activities that allow you to train certain skills.

2.1   Denken en redeneren

2.1.1   Wiskunde

Voorbeeldvraag: "Bereken 4 * 0,2="

A. Jouw positie

- Jouw score is 10.4 op 20.
- Het gemiddelde is 14.83 op 20.
- 25% van je medestudenten scoort 11.8 of minder op 20.
- 25% van je medestudenten scoort 16 of meer op 20.

■ De middelste 50% van de studiekiezers
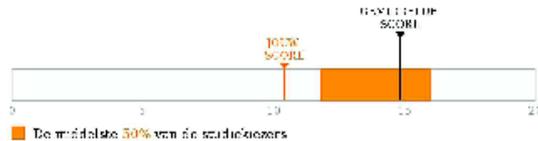
B. Jouw vaardigheid

Wat betekent jouw score? Deze test meet jouw basisvaardigheden wiskunde. Daarnaast geeft het ook een beeld van je algemene redeneervaardigheden.

Je scoort laag op de test 'basisvaardigheden wiskunde'. Schaaf deze vaardigheden zeker bij.

C. Wat nu?

2

Chapter 7: General Discussion

When making a study choice, prospective students require three fundamental pieces of information: (1) a clear understanding of the *self* (abilities, interests); (2) knowledge of the requirements of the *environment* (conditions of success, advantages and disadvantages); (3) true reasoning on the *relations of these two* groups of facts (Parsons, 1909, as cited in Brown, 2002, p.5). As a result, career choice theorist emphasize that an optimal career choice process should rely on the exploration of both the self and the environment. Prospective students should begin with a broad exploration of talents and interests, continue with the crystallization of a narrower set of specific career options, and eventually make concrete choices about jobs and careers (Feldman & Whitcomb, 2005). The quality of this study choice process has been shown important for subsequent academic outcomes (see e.g., Germeijs & Verschueren, 2007). Yet, findings on how prospective students actually accomplish their choice are discouraging. Students have a lack of knowledge about themselves and about the environment when choosing a higher education study program (Grotevant & Durrett, 1980; Wessel et al., 2008) and Flemish students only spend a limited amount of time on the exploration of their study option (Van Daal et al., 2013). Several authors (McGrath et al., 2014; Oppedisano, 2009; Vossensteyn et al., 2015) have suggested that supplying accurate information prior to enrolment improves the ability to select suitable study routes.

Therefore, we set forth to construct an instrument that delivers such information in order to facilitate decision processes in prospective students. In doing so, we followed the rationale of study choice theory by focusing on the match between personal attributes and the characteristics of the study program environment. These personal attributes typically encompass the combination of vocational interests and specific abilities (Pässler & Hell, 2012). Thus, during the construction of our instrument, we developed both a module that allows students to match their personal interests to study programs and an assessment of personal

183

competencies and their match with programs. Additionally we focused on another main component: the feedback that activates students in order to support study achievement and retention in higher education.

Research overview

Interests

Chapter 3 described the development and initial validation of SIMON-I, a new and freely available interest measure tailored to a target audience of students on the verge of selecting a higher education study program in Flanders. Overall, findings confirmed the validity and usefulness of SIMON-I and its feedback tool in the context of educational guidance and counseling.

SIMON-I was based on John Hollands' RIASEC model (Holland, 1997). The structural validity of the measure was demonstrated. Both confirmatory factor analysis (CFA) (Browne's Covariance structure modelling approach, Browne, 1992) and randomization test of hypothesized order relations (RTOR) (Hubert & Arabie, 1987; Tracey & Rounds, 1993) confirmed the underlying circumplex structure. The results of RTOR (CI = .95) surpassed common U.S. benchmarks (CI = .48) and values that were previously established with a translation of Holland's Self-Directed Search in a Flemish population of higher education students (CI = .69, Wille et al., 2014). This could suggest that SIMON-I is better at capturing the circular order of interests compared to the SDS in Flemish higher education students.

SIMON-I also contains a new academic-track-scale that deals with the often difficult choice between academic (tertiary-type A) versus vocational (tertiary-type B) programs. The rationale behind this scale was that entering an academic track requires sufficient interest in study activities that are typical for all study programs at this level. Analyses showed mean differences in scores on the 'Academic' scale between students enrolled in academic programs ($M = 54.24$, $SD = 30.05$) and those in vocational programs ($M = 40.06$, $SD = 27.55$) ($t(3960) =$

8.40, $p < .001$, Cohen's $d = .49$). The Academic scale captured the preference for academic study activities, irrespective of a specific field of interest.

In sum, internal consistencies of the subscales were good (ranging from .83 to .93) and the underlying circumplex structure of the scales was confirmed. Furthermore, gender differences in scale scores paralleled those found in previous research (Su et al., 2009). The newly developed 'Academic' scale allowed for differentiation between academic and vocational interests, both across and within fields of study. As for the output, results of one-way multivariate analysis of variance (MANOVA) confirmed that respondents from different study programs scored significantly higher on the interest scale that corresponds to the theoretical position of their study program on the circumplex. Finally, the majority of respondents (91.2%) tended to agree with the interest profile and with the corresponding programs they received (77.4%).

Evidence from chapter 2 showed that SIMON-I indeed allows identification of personally relevant study program options. When applying the matching algorithm to a sample of successful students ($N = 4,227$), 86.2% of them would receive their own study program as a suggestion in the SIMON-I output, which is higher than what has been found using other instruments. Harrington (2006), for example, found that 76% of the students graduated a major congruent with their Career decision-making system (CDM, Harrington & O'Shea, 1993) scores. In addition, congruence analyses using the C-index (Brown & Gore, 1994), which compares RIASEC codes based on the hexagonal distance between the letters, showed that the agreement between individual codes and study program codes was significantly higher than the mean of 9 ($C = 14.34$, $t = 110.92$, $p < .001$). In comparison, Wessel et al. (2008) found a mean correspondence of 10.48 (SD = 3.63) between students' interests and college major using the Strong Interest Inventory. These correspondence analyses support that SIMON-I allows for identification of personally relevant study program options.

185

When evaluating the use of SIMON-I, 55.8% of a sample of secondary education students ($N = 315$) said SIMON-I helped them in their choice process and 55.4% indicated that SIMON-I encouraged them to look into study options they had never even considered before. These numbers demonstrate that SIMON-I does aid study program choice and the (in-depth) exploration of options.

Therefore, it was concluded that SIMON-I was a promising tool that encourages the exploration of study options when making a vocational choice, be it academic or more vocationally oriented.

Competencies

A second component of this dissertation concerned the assessment of personal skills and abilities and their relation with academic performance in specific study programs.

Chapter 4 discussed the construction of a test that assesses basic mathematical skills considered vital to pass introductory statistics courses. With only 20 items, the test is very brief and thus quick and easy to administer. Results showed that the score on this test significantly contributed to the prediction of passing the statistics course, adding 8% of the explained variance over and above secondary education qualification and hours of mathematics instruction. Moreover, the mathematics test score also explained 4.2% of the variance in overall academic achievement in psychology and educational science programs. Still, although this mathematics test appeared to be valuable, its predictive validity in other programs was not yet demonstrated. Moreover, to provide (prospective) students with valid success expectations in higher education programs a broader range of variables needed to be taken into account.

These issues were taken up in chapter 5. A first proposition concerned the predictive validity of mathematical skills, other cognitive skills and educational background for academic achievement, as well as the incremental validity of personality, metacognitive, self-efficacy and

motivational factors. Secondly, we investigated whether the predictive power of variables varied across academic study programs.

As hypothesized, for most academic disciplines cognitive predictors and background factors as well as non-cognitive predictors (especially conscientiousness and self-efficacy/motivation/test anxiety) explained a significant part of the variance in academic achievement. On average, 7% of the variance in GPA was accounted for by educational background (hours of mathematics instruction in secondary education) and 8% by cognitive factors (basic mathematical knowledge, vocabulary knowledge and reading comprehension). Conscientiousness added on average 3% to the explained variance and motivational variables (controlled and autonomous motivation and self-efficacy) and test anxiety an additional 7%. Interestingly, two self-efficacy dimensions could be distinguished that predicted achievement differently. The 'comprehension' dimension had negative effects, whereas the dimension of 'effort' showed a positive relation with academic achievement. Contrary to expectations, metacognition did not predict academic achievement in any of the study programs.

The cognitive factors alone explained on average 8% of the variation in GPA, which is slightly lower than the 12% that is generally found (Kuncel & Hezlett, 2010). Also, the highest correlation between the included cognitive measures and GPA was .33 (mathematics test score in the faculty of Law). This is slightly lower than meta-analytic estimates of correlations between admissions tests and GPA, which range between .35 and .46 (Kuncel & Hezlett, 2007). Note however that this lower correlation is an artefact of the nature and goal of SIMON. As SIMON focusses on the identification of (prospective) students who lack *basic* abilities, the cognitive tests are easy for many respondents. As a result, the tests do not reflect differences in performance at the high end of the performance rage. Thus, there is a ceiling effect that limits the size of the correlations.

Taken together, the predictors explained on average 23% of the variance in GPA and passing, which corresponds to what was previously found in literature (see e.g., Robbins et al., 2004). The increase in explained variance of non-cognitive variables supports their inclusion for orientation and admission decision.

The nature of the relations between variables was also investigated using a path model based on the control-value theory of achievement emotions (Pekrun, 2006). This theory provides an integrative framework for analyzing the antecedents and effects of emotions experienced in achievement and academic settings. More specifically, the model predicts that achievement emotion (in this case test anxiety) is influenced by motivation and by self-efficacy and that all of these, combined with cognitive and metacognitive factors, affect performance. In general, support for the model was found, with all variables (except metacognition) predicting GPA. Also, achievement emotion was indeed influenced by motivation and by self-efficacy. While the control-value theory of achievement emotions discusses many antecedents of academic emotions and performance, it does not include big five personality factors. As we found that test anxiety was also affected by conscientiousness, we extended the Pekrun model.

Study results also showed that the predictive validity of non-cognitive factors for academic performance varied across study programs. Prediction accuracy was improved when taking into account the specific study program. As SIMON-C focuses on a high prediction accuracy in the high-risk group, thereby limiting false negative advice, a sensitivity of 95% was chosen as acceptable, which corresponds to a maximum of 5% of at-risk students that would unjustly get a warning that their studies are difficult to attain. Using this 95% sensitivity to select the corresponding cut-offs, 3.7% of the failing students were identified as at-risk based on the total sample prediction. In contrast, by using program specific predictions, this percentage rose to 13.4. Thus, using a program-specific prediction more students could be correctly identified as at-risk.

This second proposition of variability of predictive power across programs was further confirmed by the fact that the path model of relations between variables showed structural variance. This indicates that prediction of academic achievement would be optimized when relying on program-specific predictors or predictor levels.

Further support for the criterion validity of SIMON-C was described in chapter 2. The available data of newly enrolling students at Ghent University ($N = 8,653$) showed that 10% received a low chance of passing for the program they enrolled in. Eventually, only 6% of them passed their first year. These students obtained on average 41% of their ECTS credits, which corresponds to a chance of attaining the degree in 4 years (timely graduation of 3 years + 1 extra year) of 1.5%. In comparison, 70% of the students with a high SIMON-C chance actually passed. These students obtained on average 87% of their ECTS credits, which corresponds to a chance of attaining the degree in 4 years of 85%.

Feedback

A third important aspect in the construction of SIMON involved the consequential validity of the instrument, which is a part of construct validity that refers to the consequences of test interpretation and use (Messick, 1995). In light of its aim to increase achievement and retention by giving career choice advice, it was important to know whether and how such information had an effect on career goal management. Some students react to negative attainability feedback by doubling their efforts (the goal management strategy of assimilation; Brandtstädter & Rothermund, 2002) whilst for others such feedback leads them to abandon the goal (the strategy of accommodation; Brandtstädter & Rothermund, 2002). In the context of SIMON it was important to know whether feedback does have an effect, and if so, what kind of effect, and whether this effect is influenced by personal characteristics of the respondent.

This topic was tackled in Chapter 6. We examined whether receiving negative attainability feedback affected the career goal management strategies of assimilation and

accommodation. We further investigated whether expectancy-value variables and the perception of feedback mediated effects on career goal management.

Results showed that giving negative attainability feedback encouraged students to put more effort into their studies (assimilation, standardized coefficient = .29, $p < .001$), but even more so, it prompted them to explore other options that might be more viable for them (accommodation, standardized coefficient = .42 $p < .001$). Yet, the latter effect was somewhat undermined by the perceived accuracy of the feedback. Students who received negative feedback, depreciated its relevance and credibility. In turn, when feedback was perceived to be less accurate, this had negative effects on both accommodation and assimilation. Thus, although the direct positive effect was larger, there was also an indirect negative effect of receiving negative feedback on goal management strategies, which was mediated by perceived accuracy of feedback.

As for personal characteristics, lower motivation led to lower assimilation and to higher accommodation, as hypothesized. Yet, although it was expected that highly self-efficacious students would be more likely to persist (Williams, Donovan, & Dodge, 2000), we found no significant effect of self-efficacy on accommodation or assimilation.

The main conclusion was that giving negative attainability feedback did activate over two thirds (68.5%) of the students early on in their academic trajectory. Some doubled their efforts and others considered changing study programs. Both of these feedback effects are deemed beneficial for students' academic trajectories.

In sum, we described the construction of a valid tool for assessing the match between a (prospective) students' interest and specific study programs (chapter 3). In chapter 4, we confirmed that knowledge of basic math operations is vital to academic success in the first year at university. In chapter 5, we successfully extended this notion to other competencies by showing that the assessment of basic skills and abilities allow the identification of students that

have a high probability of failing their first year of higher education. And in chapter 6 we showed that giving feedback helps in alerting and activating individuals with potential deficits.

Strengths and Implications

One of the major advantages of SIMON is that it focusses on the lower end of the ability range. Because of the very basic level of the assessed skills and abilities the main focus is on growth potential and not on the knowledge that may or may not have been gained through secondary education. As students with a lower socio-economic status (SES) are disproportionally more often enrolled in secondary education tracks that prepare less for higher education (Groenez, Van den Brande, & Nicaise, 2003), the focus on the identification of a lack in very basic prerequisites is potentially a powerful aid in promoting social fairness. According to Müller (2014), there are three main factors that impede adolescents with a lower SES from choosing a higher education pathway, even if they are successful in school. A first is the cost of attending higher education. The second is the 'status-maintenance motive' which means that families want their children to attain a status at least as high as their own. Children with a lower SES can achieve a status equivalent to that of their parents with less investment in education. As a consequence, they are less inclined to attend higher education. And the third obstacle is that they tend to have a pessimistic view on the likelihood that their children will succeed in higher education. Indeed, outcome expectations tend to be lower in low SES youth and this plays an important role in career decision making. Thompson and Subich (2006) for example, found a mediation effect suggesting that any causal effect of social status on career certainty may occur through the mechanism of self-efficacy. SIMON can explicitly play a role in lifting this last barrier by providing realistic information and by strengthening self-confidence. Diemer, Wang, and Smith (2010) for example, found that clarifying vocational interests by using interest inventories can help low SES youth to select congruent educational environments. Prospective students may also unexpectedly receive a high chance of success which can

encourage them to attain higher education when they were first ruling this option out. Accordingly, SIMON can play an important part in reducing social inequalities in higher education.

Our data also showed (chapter 2) that SIMON more often correctly identified low (17.4%) than other SES (14.3%) risk students. Thus, low SES risk students were more often stimulated to reconsider their program choice or to participate in remedial activities than other risk students. This too may be a leverage for social equality in higher education. For example, when students switch to a more attainable study program early in the academic trajectory they can be guarded from the high financial and motivational cost of failure. Also, by participating in remedial activities that were designed specifically to nurture academic skills and abilities, students can advance their chances of success.

The fact that SIMON includes non-cognitive factors also has important advantages. Not only do they increase prediction accuracy, but they also provide prospective students with a broader view of their personal capacities. This is important as non-cognitive factors can compensate a lack of cognitive or test-taking ability (Komarraju, Ramsey, & Rinella, 2013). In line with Kyllonen (2012) and Allen (1999) it is argued that it is indispensable to include non-cognitive variables for orientation and admission decisions. Although the use of self-evaluated non-cognitive variables could be problematic because of social desirability issues, this problem is far less manifest when the test is used for study orientation instead of selection purposes, as is the case in Flanders. Still, it has been argued that non-cognitive variables may increase student success even in selective environments, for example when used post-enrollment for assisting high-risk students (Allen, Robins, & Sawyer, 2009) which parallels the application of SIMON post-enrollment. Moreover, in chapter 5 we showed that online self-evaluation questionnaires can be valid assessments of non-cognitive variables. As such, they can be a valuable and cheap alternative to labor-intensive selection procedures such as carrying out

interviews and screening letters of recommendation or essays. This use is of course conditional on whether their validity holds in selective environments, which should be tested in future research. In any case, the inclusion of non-cognitive variables also opens up possibilities for the design of interventions. As these factors are important to succeed in higher education, designing remedial activities to boost these skills could enhance academic performance.

The differential predictive validity of specific cognitive and non-cognitive factors across study programs also bears important implications for both research and counseling. In light of the findings of differential predictive validity across study programs, researchers should be cautious when generalizing findings from specific samples to all other study programs. We recommend investigators to replicate findings across student populations and to at least mention the specific study program their subjects were taken from and to limit their conclusions to this program. As for career counseling, it was shown that it is possible and valuable to make context specific predictions even when using a uniform test battery. This is especially helpful to prospective students because they can access one single platform to get information on their match with multiple programs. This is a huge benefit over taking several program-specific tests as these require the prospect to already have an idea of which programs they want to evaluate in light of their skills and abilities. Testing through one integrated platform thus increases study exploration because users may get a match with a study program they had never considered before. Moreover, it was demonstrated that program-specific study advice can be more accurate than general predictions, which is of course very valuable.

In conclusion, we demonstrated that SIMON-I and SIMON-C have important practical implications. Above all, these components resulted in a valid tool that is ready for practical use, which was the aim of this dissertation. SIMON assesses the most relevant personal attributes that have been related to academic achievement in higher education and it links these to specific study program environments. After completing SIMON-I, users receive a list of study programs

that correspond to their personal interest profile. This allows them to broaden their perspective and to look into options they may never had considered before. When filling out SIMON-C, prospective students get a clear and validated chance of success. The accuracy of this predicted probability of passing is very high (95%) in the target group of users who have a very high risk of failure in their first year of higher education. The instrument was designed to be used both for a broad exploration of personal interests and competencies and of the study program environments as well as for in-depth exploration of a specific set of career options. As such, it is applicable in any phase of the study choice process. SIMON can even be used post-enrollment, to identify and activate enrolled students who have a high risk of failure.

The use of instruments that assess personal interests and abilities and the fit with educational programs such as SIMON can be an important leverage to increase student retention. At the very least, they enable an informed choice. In a system with open access to virtually all majors, this should encourage students to choose a program that maximizes their chance of success. Furthermore, although many studies have shown that feedback does not always lead to the desired effects (see e.g., Kluger & DeNisi, 1996), giving negative attainability feedback can and does activate enrolled students to double their efforts or to reconsider their choice of study program (chapter 6), which is expected to further enhance student success.

Limitations and directions for future research

General limitations and future directions

A first opportunity for future research concerns the adoption of a more longitudinal approach to the work on both interests and capacities as well as on feedback. With regards to interests, secondary education samples could be followed longitudinally to further investigate the validity of SIMON-I for predicting study program choice and performance results. The same applies to newly enrolled students at the start of their higher education trajectory. From a

person-environment fit perspective, students who are enrolled in a program that suits their interests (as assessed by SIMON-I) should persist more and perform better than students in a program that does not fit. Examining this requires tracking new incoming students until they graduate. SIMON-C would also benefit from a more longitudinal approach. As for now, only first year academic success was predicted. And although the first year is a powerful indicator of further academic performance (Allen, 1999; de Koning et al., 2012), follow-up studies should examine how the results generalize to long term performance indicators such as timely graduation. The same applies to the results of the study on feedback: it is not possible yet to evaluate whether the attainability feedback will have long term effects on assimilative and accommodative processes. It is possible that receiving initial negative attainability feedback amplifies the effect of receiving disappointing exam results. As such, it might have more impact on career goal disengagement than we now presume. Yet, such longer term effects require a longitudinal follow-up of students.

Second, the interest and the competence assessments are separate entities in the current version of the SIMON instrument. While the competencies segment is focussed on the prediction of achievement, the interests part now solely presents what programs suit the respondent. Yet interests, and especially the fit between a person and his or her environment is also predictive of academic success (Nye et al., 2012). Therefore, it may be justified to integrate the interests and competencies components of SIMON. By including the level of congruence between a persons' interests and a study program in the prediction of program-specific chances of academic success, prediction accuracy may be further improved. As such, this could provide the prospective students with a more integrated image of their match with specific programs.

A third important general direction for future research pertains to the issue of fairness. Although SIMON is not a selection tool and its advice does not have consequences for admission, it is highly important that the instrument does not exhibit any bias towards specific

groups. Given the societal goal of broad access to educational opportunities, understanding the effect of SES is of utmost importance (Sackett et al., 2012). Currently, there is no evidence that SIMON (dis)advantages any group based on SES, gender or language background (see chapter 2). Yet, in the further development of the instrument it should be continuously monitored that the test results are not influenced by SES and do not disadvantage any underrepresented groups.

Interests: future directions

A number of limitations pertain specifically to the interests component of this dissertation. A first important issue concerns gender fairness in interest inventories (Pässler, Beinicke, & Hell, 2014). As men and women are consistently found to differ in vocational interests, with men scoring higher on Realistic and Investigative interests and women favoring Artistic, Social and Conventional activities and occupations (see e.g., Su et al., 2009), some have argued that these differences are a result of gender bias in interest inventories. One source of bias could be differential item functioning (DIF), which occurs when men and women show differing probabilities of endorsing items after matching on the underlying trait that the item is intended to measure (Zumbo, 1999, p.12). This would imply that men and women would receive a different set of suggested study programs even though they have the same level of interests. To rule this type of gender bias out we applied a IRT-based procedure, SIBTEST (Shealy & Stout, 1993), to the SIMON-I items. Results of these tests were described in chapter 3 and showed that 51% of the items showed bias. Importantly, with 47.6% of the bias items favoring women and 52.4% favoring men, there was no systematic bias against men or woman in any of the scales.

Still, we remain wary of confirming gender differences as a result of gender-biased interest scales. Therefore, gender-fairness should be continuously monitored by analyzing data from additional samples to assess the need to replace items in order to obtain more gender-fair

interest scales. Continued data gathering and analyses are also essential to the further examination of the psychometric properties of SIMON-I.

Second, more work is needed on the Academic scale. We observed genuine score differences between students in the academic and the vocational track, but the generalization of these differences to all fields of study is subject to further scrutiny.

Finally, as described in chapter 3, the profiling of study programs needs more attention. We initially departed from an expert judgment method, by which vocational interest model experts generated RIASEC profiles of study programs. Still, as the dataset is progressively expanding (see chapter 1), an 'incumbent method' (Rounds et al., 1999) can be incorporated to assign Holland codes. This implies the use of the empirically established scores per program to refine the profiles generated by experts. Relatedly, special attention can be devoted to the algorithm that matches a persons' interests profile to the study program environments. In the current work, the congruence between these two is based on the match between the RIASEC letter codes. This procedure typically checks the correspondence between the top letters (e.g., the top three) of the person and environment Holland code. Although this method is widely adopted, it has been rightfully criticized for not incorporating the entire interest profile (Tracey, Allen, & Robbins, 2012). The application of more recent correlational approaches to person-environment fit (see e.g., Allen & Robbins, 2010; Tracey et al., 2012) seems especially promising to counter this obstacle and has recently been applied to SIMON data (Schelfhout et al., 2017).

Comptencies: future directions

This dissertation also addressed the prediction of academic success through cognitive and non-cognitive factors. Although 23% of variance in academic achievement was accounted for, a lot remains unexplained. This leaves a plethora of options for further research as the predictive validity of many other variables can be examined on top of the ones that were already

included in this dissertation. Notable examples are self-control (see e.g., Tangney, Baumeister, & Boone, 2004), the utility value of the course (Eccles & Wigfield, 2002), personality constructs other than conscientiousness (such as openness to experiences), academic emotions such as boredom and enjoyment (Detmers et al., 2011; Pekrun et al., 2002), and many others. In fact, it is worth examining the incremental predictive validity of all factors that have somehow been related to academic achievement as their inclusion may allow a better prediction and thus a more comprehensive model of academic achievement.

The differential predictive power of variables for specific study programs is also potentially a very rich area of research. Future studies should replicate and extend our findings. For example, STEM (Science, technology, engineering and mathematics) programs were not included in chapter 5. Also, a deeper understanding of why there might be differences in what predicts college success across programs is up for further discussion. In this context it may be worthwhile to take up the concept of multidimensionality of student performance (Kuncel et al., 2001; Oswald et al., 2004). These dimensions include intellectual behaviors (such as knowledge and learning), interpersonal behaviors (such as leadership), and intrapersonal behaviors (such as adaptability and perseverance) (Oswald et al., 2004). It would be interesting to thoroughly examine if, how and why the explanatory value of these dimensions varies by major.

In chapter 5, we found that two dimensions of self-efficacy could be distinguished: one called 'effort' (the confidence one has that one will put in the effort to succeed) and another one labeled 'comprehension' (the confidence one has that one will understand the contents of the courses). Whilst the first had a positive relation with academic achievement, the latter had adverse effects on achievement. This may explain why some researchers have found negative effects of self-efficacy on performance (Vancouver et al., 2002; Vancouver, Thompson, & Williams, 2001). Vancouver and Kendall (2006) hypothesized that high self-efficacy can lead

to diminished effort which negatively affects performance, thus suggesting that effort mediates the relationship between self-efficacy and performance (Vancouver et al., 2002). Therefore, singling out a distinct effort dimension apart from other self-efficacy facets may improve our insight in its effects on achievement. Multon et al. (1991) also argued that self-efficacy can be too high, leading to gross overestimations of one's ability and causing one to attempt tasks well beyond one's potential. This issue is closely related to the concepts of optimism and overconfidence. Stone (1994) indeed found that high self-efficacy led to overconfidence in one's abilities which in turn lessened the attention and effort in a performance task. With regards to self-efficacy (and study program choice alike) it seems key to hold realistic expectations. Indeed, students who have realistic expectations perform better than those who do not (Nicholson, Putwain, Connors, & Hornby-Atkinson, 2013). Although optimism is generally regarded as a positive characteristic, some authors warn that unmitigated optimism can be problematic, especially in novel settings such as higher education because individuals have little relevant experience on which to base such expectations (Haynes, Ruthig, Perry, Stupnisky, & Hall, 2006). This effect has been demonstrated. For example, not only did overconfidence lead students to psychologically withdraw from the academic context (i.e., view grades as progressively less important), but their unrealistically high goals also caused them to drop out altogether (Robins & Beer, 2001). Thus, although a higher self-efficacy is generally associated with better academic achievement, it does hold a potential risk to overestimate one's abilities which may lead to a decrease in effort and may result in lower performance. Discriminating between multiple dimensions of self-efficacy may advance our understanding of its effects on academic success. Future studies need to address whether the dimensions of effort and comprehension as well as their association to achievement can be replicated in other samples. In general, more research is required on different facets of self-efficacy and on how these interact with overconfidence in regulating behavior (Lent, Brown, & Gore, 1997).

Moores and Chang (2009) provide an example of such a study. They found that overconfidence leads to a significant negative relationship between self-efficacy and subsequent performance. They showed that high self-efficacy could have a debilitating effect on performance depending on the level of overconfidence and on the extent to which self-efficacy was adjusted based on explicit performance feedback. Their study was concluded by highlighting the importance of providing timely feedback, a suggestion that is just as valuable with regards to the study choice process.

Feedback: future directions

The importance of feedback that activates the receiver can indeed not be emphasized sufficiently. Up until a decade ago, little attention had been given to the reporting of assessment results (Goodman & Hambleton, 2004). Since then, test score reporting is a topic that matters. Several works have been published that contain guidelines on the content and form of feedback reports in order to maximize the ease of interpretation (Hattie, 2009; Roberts & Gierl, 2009, 2010; Vezzu, Van Winkle, & Zapata-Rivera, 2012). Nevertheless, many feedback related topics deserve further attention.

In general, feedback studies in ecologically valid settings are scarce and very few general conclusions can be drawn (Rakoczy, Harks, Klieme, Blum, & Hochweber, 2013). Individual differences probably influence the relationship between feedback and outcomes (Shute, 2008). Yet, these differences are not yet fully understood. Also, the role of feedback after diagnostic or high-stakes assessments has received little attention (Harrison et al., 2013) and we know little about the influence of feedback on career development (Creed et al., 2015). In addition, most studies have focused on performance improvement following negative feedback. Yet, in the context of unattainable career goals it would be especially interesting to examine what feedback characteristics are important to encourage goal disengagement and what factors determine the acceptance of such feedback.

Specifically with regards to SIMON, we found that feedback affected the career goal management strategies in students who had already enrolled in a study program, but it is still unclear whether prospective students would react in similar ways to such negative attainability feedback. When investigating the effects of feedback on prospective students' choice behavior, special attention could be devoted to the optimization and personalization of feedback strategies. For example, the level of feedback specification may be relevant to consider. Carroll et al. (2009) experimentally induced a career goal to demonstrate that the process of goal disengagement is more likely when negative feedback fully specifies the implications of continued striving for an unattainable goal as opposed to when giving unspecified feedback. It may also be important to study interactions between feedback characteristics and individual differences. We tried to lift a corner of this veil by examining whether self-efficacy and motivation influenced the relation between feedback and goal (dis)engagement. Although social cognitive theories would suggest that highly self-efficacious students are more likely to persist in the presence of goal-performance discrepancies (Williams et al., 2000), there was no significant effect of self-efficacy on either accommodation or assimilation. This requires further examination. For example, only a one-item general measure of self-efficacy was used in that specific study. Maybe the inclusion of the different dimensions of self-efficacy would advance our understanding of individual differences in reactions to feedback. It is probable that those who score high on the comprehension dimension are more likely to overestimate their abilities and are thus most in need of honest attainability feedback. It is likely that this groups' responsiveness to feedback would increase when using fully specified feedback. Thus, research into this matter could prompt the further development of personalized feedback strategies.

Finally, as feedback is an essential part of goal-setting behavior, the optimization of feedback goes hand in hand with an expansion of the knowledge on study choice processes in general and individual differences in goal disengagement more specifically. The processes and

factors that influence goal disengagement are still unclear (Heckhausen, Wrosch, & Schulz, 2010). Moreover, there is limited research into disengagement from career goals during the study choice process (Creed & Blume, 2013). Although the current dissertation did contribute to the understanding of the career goal (dis)engagement process, much remains unclear with regards to what factors predict who will persist in engaging with an unattainable goal and who will disengage (Boudrenghien et al., 2012; Carroll et al., 2009; Converse, Steinhauser, & Pathak, 2010; Tomasik & Silbereisen, 2012). Together, all of these feedback and goal engagement related issues may shed further light on how to ease feedback acceptance and how feedback can promote goal disengagement in ecologically valid settings.

Conclusion

This dissertation described the construction and validation process and procedures of SIMON, a tool that allows (prospective) students to evaluate their match with specific study programs. The four included studies concerned three major components of the instrument: the match between the personal interests and the study program environment, the evaluation of personal skills and abilities in light of specific study programs and the provision of activating feedback. The research on these components has contributed to the understanding of post-secondary academic achievement and study orientation processes. But most importantly, this research led to a valid instrument that is now ready-for-use, both for prospective and for newly enrolled students.

References

Allen, D. (1999). Desire to finish college: an empirical link between motivation and persistence. *Research in Higher Education, 40*, 461-485.

Allen, J., & Robbins, S. (2010). Effects of interest-major congruence, motivation, and academic performance on timely degree attainment. *Journal of Counseling Psychology, 57*(1), 23-35. doi:10.1037/a0017267

Allen, J., Robbins, S., & Sawyer, R. (2009). Can measuring psychosocial factors promote college success? *Applied Measurement in Education, 23*(1), 1-22. doi:10.1080/08957340903423503

Boudrenghien, G., Frenay, M., & Bourgeois, E. (2012). Unattainable educational goals: Disengagement, reengagement with alternative goals, and consequences for subjective well-being. *European Review of Applied Psychology-Revue Europeenne De Psychologie Appliquee, 62*(3), 147-159. doi:10.1016/j.erap.2012.04.002

Brandtstädter, J., & Rothermund, K. (2002). The life-course dynamics of goal pursuit and goal adjustment: A two-process framework. *Developmental Review, 22*(1), 117-150. doi:10.1006/drev.2001.0539

Brown, D. (2002). Introduction to Theories of Career Development and Choice: Origins, Evolution, and Current Efforts. In D. Brown (Ed.), *Career choice and development* (Fourth ed.). San Francisco: John Wiley and Sons, Inc.

Brown, S. D., & Gore, P. A. (1994). An evaluation of interest congruence indices: Distribution characteristics and measurement properties. *Journal of Vocational Behavior, 45*(3), 310-327. doi:10.1006/jvbe.1994.1038

Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika, 57*(4), 469-497. doi:10.1007/bf02294416

Carroll, P. J., Shepperd, J. A., & Arkin, R. M. (2009). Downward Self-Revision: Erasing Possible Selves. *Social Cognition, 27*(4), 550-578. doi:10.1521/soco.2009.27.4.550

Converse, P. D., Steinhauser, E., & Pathak, J. (2010). Individual differences in reactions to goal-performance discrepancies over time. *Personality and Individual Differences, 48*(2), 138-143. doi:10.1016/j.paid.2009.09.010

Creed, P. A., & Blume, K. (2013). Compromise, Well-Being, and Action Behaviors in Young Adults in Career Transition. *Journal of Career Assessment, 21*(1), 3-19. doi:10.1177/1069072712453830

Creed, P. A., Wamelink, T., & Hu, S. (2015). Antecedents and consequences to perceived career goal-progress discrepancies. *Journal of Vocational Behavior, 87*, 43-53. doi:10.1016/j.jvb.2014.12.001

de Koning, B., Loyens, S., Rikers, E., Smeets, G., & van der Molen, H. (2012). Generation Psy: Student characteristics and academic achievement in a three-year problem-based learning bachelor program. *Learning and Individual Differences, 22*(3), 313-323. doi:10.1016/j.lindif.2012.01.003

Detmers, S., Trautwein, U., Lüdtke, O., Goetz, T., Frenzel, A. C., & Pekrun, R. (2011). Students' emotions during homework in mathematics: Testing a theoretical model of antecedents and achievement outcomes. *Contemporary Educational Psychology, 36*(1), 25-35. doi:10.1016/j.cedpsych.2010.10.001

Diemer, M. A., Wang, Q., & Smith, A. V. (2010). Vocational interests and prospective college majors among youth of color in poverty. *Journal of Career Assessment, 18*(1), 97-110. doi:10.1177/1069072709350906

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

Feldman, D. C., & Whitcomb, K. M. (2005). The effects of framing vocational choices on young adults' sets of career options. *Career Development International, 10*(1), 7-25. doi:10.1108/13620430510577600

Germeijs, V., & Verschueren, K. (2007). High school students' career decision-making process: Consequences for choice implementation in higher education. *Journal of Vocational Behavior, 70*(2), 223-241. doi:10.1016/j.jvb.2006.10.004

Goodman, D. P., & Hambleton, R. K. (2004). Student test score reports and interpretive guides: Review of current practices and suggestions for future research. *Applied Measurement in Education, 17*(2), 145-220. doi:10.1207/s15324818ame1702_3

Groenez, S., Van den Brande, I., & Nicaise, I. (2003). *Cijferboek sociale ongelijkheid in het Vlaamse onderwijs*. Retrieved from http://informatieportaalssl.be/archiefloopbanen/rapporten/LOA-rapport_10.pdf:

Grotevant, H. D., & Durrett, M. E. (1980). Occupational knowledge and career development in adolescence. *Journal of Vocational Behavior, 17*(2), 171-182. doi:10.1016/0001-8791(80)90002-0

Harrington, T. F. (2006). A 20-year follow-up of the Harrington-O'Shea Career Decision-Making System. *Measurement and Evaluation in Counseling and Development, 38*(4), 198.

Harrington, T. F., & O'Shea, A. J. (1993). *The Harrington-O'Shea career decision-making system revised. Manual.* Retrieved from Circle Pines, Minnesota:

Harrison, C. J., Konings, K. D., Molyneux, A., Schuwirth, L. W. T., Wass, V., & van der Vleuten, C. P. M. (2013). Web-based feedback after summative assessment: how do students engage? *Medical education, 47*(7), 734-744. doi:10.1111/medu.12209

Hattie, J. (2009). Visibly learning from reports: the validity of score reports. *Online Educational Research Journal.* Retrieved from http://www.oerj.org/View?action=viewPDF&paper=6

Haynes, T. L., Ruthig, J. C., Perry, R. P., Stupnisky, R. H., & Hall, N. C. (2006). Reducing the academic risks of over-optimism: The longitudinal effects of attributional retraining on cognition and achievement. *Research in Higher Education, 47*(7), 755-779.

Heckhausen, J., Wrosch, C., & Schulz, R. (2010). A motivational theory of life-span development. *Psychological review, 117*(1), 32. doi:10.1037/a0017668

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Lutz, FL: Psychological Assessment Resources Inc.

Hubert, L., & Arabie, P. (1987). Evaluating order hypotheses within proximity matrices. *Psychological Bulletin, 102*, 172-178.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254. doi:10.1.1.461.6812

Komarraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. *Learning and Individual Differences, 24*, 103-109. doi:10.1016/j.lindif.2012.12.007

Kuncel, N., Hezlett, S. A., & Ones, D. S. (2001). A Comprehensive Meta-Analysis of the Predictive Validity of the Graduate Record Examinations: Implications for Graduate Student Selection and Performance. *Psychological Bulletin, 127*(1), 162-181. doi:10.1037//q033-2909.127.1.162

Kuncel, N., & Hezlett, A. (2007). Standardized tests predict graduate students' success. *Science, 315*(5815), 1080-1081. doi:10.1126/science.1136618

Kuncel, N., & Hezlett, S. (2010). Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current Directions in Psychological Science, 19*(6), 339-345. doi:10.1177/0963721410389459

Kyllonen, P. C. (2012). The Importance of Higher Education and the Role of Noncognitive Attributes in College Success. *Pensamiento Educativo. Revista de Investigación Educacional Latinoamericana, 49*(2), 84-100.

Lent, R. W., Brown, S. D., & Gore, P. A., Jr. (1997). Discriminant and predictive validity of academic self-concept, academic self-efficacy, and mathematics-specific self-efficacy. *Journal of Counseling Psychology, 44*(3), 307-315. doi:10.1037/0022-0167.44.3.307

McGrath, C., Henham, M., Corbett, A., Durazzi, N., Frearson, M., Janta, B., . . . Schweppenstedde, D. (2014). *Higher Education Entrance Qualifications and Exams in Europe: A Comparison.* Retrieved from http://www.rand.org/pubs/research_reports/RR574

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist, 50*(9), 741.

Moores, T. T., & Chang, J. C.-J. (2009). Self-efficacy, overconfidence, and the negative effect on subsequent performance: A field study. *Information & Management, 46*(2), 69-76.

Müller, W. (2014). Educational inequality and social justice: Challenges for career guidance. *International Journal for Educational and Vocational Guidance, 14*(1), 21-33. doi:10.1007/s10775-014-9264-4

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. doi:10.1037/0022-0167.38.1.30

Nicholson, L., Putwain, D., Connors, L., & Hornby-Atkinson, P. (2013). The key to successful achievement as an undergraduate student: confidence and realistic expectations? *Studies in Higher Education, 38*(2), 285-298.

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational Interests and Performance: A Quantitative Summary of Over 60 Years of Research. *Perspectives on Psychological Science, 7*(4), 384-403. doi:10.1177/1745691612449021

Oppedisano, V. (2009). *Open University Admission Policies and Drop Out Rates in Europe*. Retrieved from

Oswald, F., Schmitt, N., Kim, B., Ramsay, L., & Gillespie, M. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89*(2), 187-207.

Pässler, K., Beinicke, A., & Hell, B. (2014). Gender-related differential validity and differential prediction in interest inventories. *Journal of Career Assessment, 22*(1), 138-152. doi:10.1177/1069072713492934

Pässler, K., & Hell, B. (2012). Do interests and cognitive abilities help explain college major choice equally well for woman and men? *Journal of Career Assessment*. doi:10.1177/1069072712450009

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315-341. doi:10.1007/s10648-006-9029-9

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91-105. doi:10.1207/S15326985EP3702_4

Rakoczy, K., Harks, B., Klieme, E., Blum, W., & Hochweber, J. (2013). Written feedback in mathematics: Mediated by students' perception, moderated by goal orientation. *Learning and Instruction, 27*, 63-73. doi:10.1016/j.learninstruc.2013.03.002

Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288. doi:10.1037/0033-2909.130.2.261

Roberts, M. R., & Gierl, M. J. (2009). *Development of a Framework for Diagnostic Score Reporting*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Roberts, M. R., & Gierl, M. J. (2010). Developing score reports for cognitive diagnostic assessments. *Educational measurement: issues and practice, 29*(3), 25-38.

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: short-term benefits and long-term costs. *Journal of Personality and Social Psychology, 80*(2), 340-352. doi:10.1037//0O22-3514.80.2.340

Rounds, J., Smith, T., Hubert, L., Lewis, P., & Rivkin, D. (1999). Development of occupational interest profiles for O* NET. *National Center for O* NET Development, Raleigh, NC*.

Sackett, P. R., Kuncel, N. R., Beatty, A. S., Rigdon, J. L., Shen, W., & Kiger, T. B. (2012). The Role of Socioeconomic Status in SAT-Grade Relationships and in College Admissions Decisions. *Psychological science, 23*(9), 1000-1007. doi:10.1177/0956797612438732

Schelfhout, S., Wille, B., Fonteyne, L., Roels, E., De Fruyt, F., & Duyck, W. (2017). *From Person-Environment Fit to Study Counseling: an Empirical Advice Set Engine*. Manuscript submitted for publication.

Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika, 58*(2), 159-194. doi:10.1007/BF02294572

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153-189.

Stone, D. N. (1994). Overconfidence in initial self-efficacy judgments - effects on decision-processes and performance. *Organizational Behavior and Human Decision Processes, 59*(3), 452-474. doi:10.1006/obhd.1994.1069

Su, R., Rounds, J., & Armstrong, P. I. (2009). Men and things, women and people: a meta-analysis of sex differences in interests. *Psychological Bulletin, 135*(6), 859-884. doi:10.1037/a0017364.

Tangney, J., Baumeister, R., & Boone, A. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271-322. doi:10.1111/j.0022-3506.2004.00263.x

Thompson, M. N., & Subich, L. M. (2006). The relation of social status to the career decision-making process. *Journal of Vocational Behavior, 69*(2), 289-301. doi:10.1016/j.jvb.2006.04.008

Tomasik, M. J., & Silbereisen, R. K. (2012). Beneficial effects of disengagement from futile struggles with occupational planning: A contextualist-motivational approach. *Developmental Psychology, 48*(6), 1785. doi:10.1037/a0026952

Tracey, T. J. G., Allen, J., & Robbins, S. B. (2012). Moderation of the relation between person–environment congruence and academic success: Environmental constraint, personal flexibility and method. *Journal of Vocational Behavior, 80*(1), 38-49. doi:10.1016/j.jvb.2011.03.005

Tracey, T. J. G., & Rounds, J. B. (1993). Evaluating Holland's and Gati's vocational-interest models: A structural meta-analysis. *Psychological Bulletin, 113*(2), 229-246. doi:10.1037/0033-2909.113.2.229

Van Daal, T., Coertjens, L., Delvaux, E., Donche, V., & Van Petegem, P. (2013). *Klaar voor hoger onderwijs of arbeidsmarkt? Longitudinaal onderzoek bij laatstejaarsleerlingen secundair onderwijs.* Antwerpen - Apeldoorn: Garant.

Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology, 91*(5), 1146. doi:10.1037/0021-9010.91.5.1146

Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology, 87*(3), 506. doi:10.1037//0021-9010.87.3.506

Vancouver, J. B., Thompson, C. M., & Williams, A. A. (2001). The changing signs in the relationships among self-efficacy, personal goals, and performance. *Journal of Applied Psychology, 86*(4), 605. doi:10.1037/0021-9010.86.4.605

Vezzu, M., Van Winkle, W., & Zapata-Rivera, D. (2012). *Designing and evaluating an interactive score report for students* (ETS RM-12-01). Retrieved from http://www.ets.org/research/contact.html

Vossensteyn, H., Stensaker, B., Kottmann, A., Hovdhaugen, E., Jongbloed, B., Wollscheid, S., . . . Cremonini, L. (2015). *Dropout and Completion in Higher Education in Europe*. Retrieved from http://ec.europa.eu/dgs/education_culture/repository/education/library/study/2015/dropout-completion-he_en.pdf

Wessel, J. L., Ryan, A. M., & Oswald, F. L. (2008). The relationship between objective and perceived fit with academic major, adaptability, and major-related outcomes. *Journal of Vocational Behavior, 72*(3), 363-376. doi:10.1016/j.jvb.2007.11.003

Wille, B., Tracey, T. J. G., Feys, M., & De Fruyt, F. (2014). A longitudinal and multi-method examination of interest–occupation congruence within and across time. *Journal of Vocational Behavior, 84*(1), 59-73. doi:10.1016/j.jvb.2013.12.001

Williams, K. J., Donovan, J. J., & Dodge, T. L. (2000). Self-regulation of performance: Goal establishment and goal revision processes in athletes. *Human Performance, 13*(2), 159-180. doi:10.1207/s15327043hup1302_3

Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and Likert-type (ordinal) item scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

Nederlandstalige samenvatting

Introductie

Toegang tot een opleiding in het hoger onderwijs in Vlaanderen is nagenoeg ongelimiteerd. Met uitzondering van de opleidingen geneeskunde, tandheelkunde en de kunsten, waar een toelatingsexamen georganiseerd wordt, kan elke student met een diploma secundair onderwijs zich inschrijven voor om het even welke opleiding hoger onderwijs. Dit systeem heeft als doel om de sociale ongelijkheid in het hoger onderwijs te minimaliseren, maar heeft als keerzijde dat het eerste jaar van het hoger onderwijs een soort 'selectiejaar' is. Slechts 40% van de eerstejaarsstudenten behaalt alle opgenomen ECTS-credits en 17% behaalt geen enkele credit (Ministerie van Onderwijs en Vorming, 2009).

De manier waarop jongeren een studiekeuze maken kan een rol spelen bij deze lage slaagcijfers. Een weldoordachte studiekeuze kan immers het studierendement verhogen. Een passende opleiding kiezen verhoogt de tevredenheid, de prestatie en de doorzetting van studenten in het hoger onderwijs (Feldman et al., 1999; Nye et al., 2012). Helaas toont onderzoek (Van Daal et al., 2013) aan dat Vlaamse jongeren weinig tijd spenderen aan het exploreren van hun opties en het maken van hun studiekeuze, zelfs amper drie maanden voor de start van het hoger onderwijs. Jongeren moeten bijgevolg gestimuleerd worden om een doordachte studiekeuze te maken, een keuze die drie fundamentele factoren behelst: (1) Een duidelijk begrip van de eigen competenties en interesses; (2) Kennis van de vereisten van de opleidingen; en (3) Reflectie op de relatie tussen deze twee voorgaande factoren (Brown, 2002).

Deze reflectie impliceert dat studiekiezers over de noodzakelijke kennis en informatie beschikken. Het is dan ook aangewezen om hen adequate handvaten aan te reiken om hun exploratie- en reflectieproces te ondersteunen. Tot op heden hadden toekomstige studenten in Vlaanderen echter geen valide instrumenten die hen stimuleren om zichzelf, de hoger

onderwijsomgeving en de afstemming tussen deze twee te exploreren en op die manier een geïnformeerde studiekeuze te maken.

Deze doctoraatsscriptie is gericht op het beschrijven van de ontwikkeling en validering van een dergelijk online assessment-instrument (SIMON (Studievaardigheden- en InteresseMONitor)). Door valide feedback te geven over persoonlijke kenmerken (competenties en interesses) en de afstemming van deze kenmerken op opleidingsmogelijkheden in het hoger onderwijs in Vlaanderen wordt een optimale studiekeuze ondersteund en gestimuleerd.

Het instrument werd gecentreerd rond twee grote componenten die corresponderen met de vragen die studiekiezers zich stellen: (1) wat wil ik studeren? en (2) kan ik slagen in deze opleiding?. De eerste vraag betreft de interesses van de studiekiezer en de afstemming hiervan op specifieke opleidingen. Omdat er een gebrek was aan valide instrumenten die de interesses van (toekomstige) studenten koppelen aan opleidingen in Vlaanderen werd er in dit onderzoek eerst gefocust op de ontwikkeling van een context-specifiek interesse-instrument met bijhorende feedbackmodule (SIMON-I). Gezien er in Vlaanderen relatief veel studenten inschrijven in het professioneel (versus het academisch) hoger onderwijs (OECD, 2014) was deze contextspecificiteit voornamelijk gericht op het discrimineren tussen de interesse in meer academische hetzij in meer professionele opleidingen.

De tweede SIMON-component behelsde de overeenstemming tussen persoonlijke vaardigheden en de vereisten van studieprogramma's. SIMON focuste hierbij op predictieve validiteit: het doel was om studiekiezers informatie te geven over hun kansen op studiesucces in een specifieke opleiding. Hoewel dit doel parallellen vertoont met dat van toelatingsproeven, verschilt SIMON fundamenteel van dergelijke testen. Terwijl selectie-instrumenten doorgaans beperkt blijven tot het testen van cognitieve factoren, werd er bij de voorspelling van studiesucces in SIMON ook rekening gehouden met niet-cognitieve factoren waarvan de

predictieve waarde eerder werd aangetoond (Credé & Kuncel, 2008; Robbins et al., 2004). Een ander onderscheid is dat SIMON focust op het testen van *basis*vaardigheden en op een grote accuraatheid van de voorspelling. Er werd gemikt op het identificeren van studenten die met een zeer grote waarschijnlijkheid niet beschikken over de noodzakelijke vaardigheden om het eerste jaar van een opleiding met succes af te ronden. Een *kleine groep studenten* moest het signaal krijgen dat een opleiding onhaalbaar is, maar deze voorspelling moest *bijzonder accuraat* zijn. Studenten die kunnen slagen moesten het voordeel van de twijfel krijgen en niet ontmoedigd worden. Daarom mikten we op een accuraatheid van 95% in de groep met lage slaagkansen. Met andere woorden: wie een negatief advies krijgt, zal in 95% van de gevallen effectief niet slagen. Een derde belangrijk verschil met toelatingsproeven is dat de resultaten van SIMON niet bindend zijn, maar vooral tot doel hebben om een optimale maar vrije studiekeuze te ondersteunen.

Om deze voorspelling van studiesucces te valideren werden nieuw ingestroomde studenten gevolgd. Zij werden bij de start van hun eerste academiejaar, wanneer hun profiel nog sterk aansluit bij dat van de studiekiezer, getest op verschillende cognitieve en niet-cognitieve vaardigheden. Vervolgens werden deze testscores gekoppeld aan hun examenresultaten op het einde van het academiejaar. Deze methode had als bijkomend voordeel dat er waardevolle informatie over de ingeschreven studenten beschikbaar is. Hoewel de primaire doelgroep van SIMON de studiekiezers zijn, bood de beschikbaarheid van valideringsdata ook voordelen voor ingeschreven studenten. SIMON kan deze laatsten immers een beeld geven van hun startpositie in het hoger onderwijs. Wanneer een ingeschreven student een hoge waarschijnlijkheid heeft om te falen in het eerste jaar kan deze student geactiveerd worden om deel te nemen aan georganiseerde remediëringsactiviteiten. Op die manier kan SIMON ook ná inschrijving het studierendement in het hoger onderwijs verhogen.

Dit brengt ons bij de derde en laatste belangrijke component van SIMON: feedback. Een instrument mag dan wel waardevolle informatie geven over de afstemming tussen persoonlijke kenmerken en opleidingen, de vraag blijft of (toekomstige) studenten geactiveerd worden door deze feedback. Als gebruikers niets doen met de feedback die zij krijgen zal SIMON het studiekeuzeproces niet ondersteunen en dus ook het succes in het hoger onderwijs niet stimuleren. In deze scriptie werd bijgevolg ook gefocust op dit belangrijke element in het valideringsproces (Duijnhouwer et al., 2012; Hattie, 2009; Messick, 1990): de evaluatie van hoe testresultaten gebruikt worden.

Deze drie componenten (interesses, competenties en feedback) vormden het onderwerp van deze doctoraatsscriptie en van het studiekeuze-instrument SIMON.

Studies in dit doctoraatsproefschrift

Interesses

In hoofdstuk 3 werd de ontwikkeling van de interessevragenlijst, SIMON-I, beschreven. Deze vragenlijst is gebaseerd op het RIASEC-model van John Holland (1997), het meest geciteerde en gevalideerde psychologisch model van beroeps- en opleidingsinteresses. Dit model karakteriseert mensen, beroepen en opleidingen aan de hand van zes dimensies die onderling kunnen gecombineerd worden, en die zo de variatie in het Vlaamse onderwijsaanbod omvatten: het Praktische, het Analytische, het Kunstzinnige, het Sociale, het Ondernemende en het Conventionele (de vertaling van de RIASEC-dimensies werden overgenomen uit Dingemanse, Van Amstel, De Fruyt, & Wille, 2009). Bovenop deze dimensies voegden we een subschaal toe die de specifieke interesse in hetzij academisch gerichte, hetzij professioneel gerichte opleidingen test.

De structurele validiteit van het onderliggende model werd op twee verschillende wijzen bevestigd (door confirmatorische factor analyse, Browne, 1992; en door randomization test of hypothesized order relations, Tracey & Rounds, 1993). Resultaten toonden ook dat de nieuwe

'academische interesse'-subschaal de interesse in academische studie-activiteiten kon vatten en dit onafhankelijk van het specifieke interesseveld. Met andere woorden: ook binnen de verschillende PAKSOC-dimensies discrimineerde SIMON-I tussen academische en professionele interesse.

De evaluatie van respondenten was bemoedigend: de meerderheid verklaarde zich akkoord met het verkregen interesseprofiel en met de daaraan gekoppelde opleidingen. Over het algemeen bevestigden de resultaten de validiteit en bruikbaarheid van SIMON-I in de context van studie-oriëntering en -begeleiding.

Competenties

In hoofdstuk 4 werd de validiteit beschreven van een test die nagaat of studenten over de noodzakelijke basisvaardigheden wiskunde beschikken om te slagen in een inleidend opleidingsonderdeel statistiek. Uit de resultaten bleek dat deze korte test (20 items) significant bijdroeg aan de voorspelling van slagen voor het opleidingsonderdeel statistiek en dit bovenop traditioneel sterke voospellers als het diploma secundair onderwijs en het aantal uren wiskundeonderwijs in het secundair. De wiskundetest voorspelde echter niet enkel het slagen voor statistiek, maar verklaarde ook 4.2% van de variantie in het slagen voor alle vakken van het eerste jaar van de opleidingen psychologie en pedagogische wetenschappen.

Hoewel waardevol, was deze test te beperkt om (toekomstige) studenten een valide beeld te geven van hun slaagkansen in het hoger onderwijs. Ten eerste moest de predictieve validiteit van de test in andere programma's dan de psychologie en pedagogische wetenschappen nog aangetoond worden. Daarnaast was de score op een wiskundetest ontoereikend om een globaal beeld van de succesverwachtingen in opleidingen te schetsen. Het in rekening brengen van een bredere waaier aan variabelen én aan opleidingen zou dan ook toelaten om ruimere en meer valide voorspellingen te doen.

Deze beperkingen werden aangekaart in hoofdstuk 5. Een eerste focus van deze studie lag op de incrementele validiteit voor het voorspellen van academisch succes van de wiskundetest, maar ook van andere cognitieve variabelen, achtergrondkenmerken, persoonlijkheid, metacognitieve vaardigheden, zelf-effectiviteit en motivationele factoren. Daarnaast onderzochten we of de voorspellende waarde van deze variabelen varieerde tussen de verschillende opleidingen.

Zoals verondersteld werd slagen het best voorspeld door een combinatie van cognitieve en niet-cognitieve variabelen (consciëntieusheid en zelf-effectiviteit/motivatie/testangst). Samen verklaarden de opgenomen variabelen gemiddeld 23% van de variantie in academisch succes, wat overeenstemt met wat eerder in de literatuur werd aangetroffen (zie bijvoorbeeld Robbins et al., 2004).

Ook de tweede hypothese werd bevestigd: de incrementele predictieve validiteit van de variabelen varieerde tussen verschillende opleidingen. Bij opleidingsspecifieke voorspellingen werden 9.7% meer falende studenten geïdentificeerd als risicostudent dan bij algemene voorspellingen (over opleidingen heen) werden gemaakt. Deze resultaten demonstreren dat de voorspelling van academisch succes kan geoptimaliseerd worden door predictoren of niveaus te laten variëren afhankelijk van de specifieke opleidingen.

Feedback

Een derde belangrijk aspect in de ontwikkeling van SIMON betrof de consequentiële validiteit (Messick, 1990) van het instrument. Als het doel van SIMON eruit bestaat het studiesucces te verhogen door studiekeuze-advies te verstrekken, moest nagegaan worden of dergelijke advies een effect heeft op het keuze- of studiegedrag van (toekomstige) studenten.

Wanneer studenten feedback krijgen dat de opleiding die ze willen volgen moeilijk haalbaar is reageren studenten op verschillende manieren. Sommigen zien dit als een stimulans om hun inspanningen te verdubbelen om alsnog te slagen (de assimilatie-strategie;

Brandtstädter & Rothermund, 2002) terwijl anderen ervoor kiezen om hun doel op te geven en een andere opleiding te overwegen (de accommodatie-strategie; Brandtstädter & Rothermund, 2002). In de context van SIMON was het belangrijk te weten of feedback een effect heeft en indien dit het geval is, welk effect en of dit effect wordt beïnvloed door persoonlijke kenmerken van de respondent. Dit onderwerp werd geëxploreerd in hoofdstuk 6.

We onderzochten of het krijgen van een lage slaagkans voor een opleiding een invloed had op de doelstrategieën assimilatie en accommodatie. We bekeken tevens of "expectancy-value"-variabelen (Wigfield & Eccles, 2000) en de perceptie van feedback de effecten op de doelstrategieën medieerden.

De resultaten toonden dat het geven van negatieve haalbaarheidsfeedback studenten aanmoedigde om meer moeite te doen voor hun studies (assimilatie), maar meer nog, het zette hen ertoe aan andere opleidingen te exploreren die meer geschikt zouden kunnen zijn (accommodatie). Dit laatste effect werd echter enigszins ondermijnd door de gepercipieerde accuraatheid van de feedback. Studenten die negatieve feedback kregen, twijfelden sterker aan de relevantie en geloofwaardigheid ervan. Wanneer feedback als minder accuraat werd beschouwd had dit negatieve effecten op zowel assimilatie als op accommodatie. Hoewel de directe positieve effecten groter waren, was er een indirect negatief effect (gemedieerd door de gepercipieerde accuraatheid) van negatieve feedback op doelstrategieën.

Zoals verondersteld leidde lagere motivatie tot lagere assimilatie en tot hogere accommodatie. Hoewel het de verwachting was dat studenten met hogere zelf-effectiviteit een grotere mate van assimilatie zouden vertonen (Williams et al., 2000), vonden we geen dergelijk effect.

De belangrijkste conclusie was dat het geven van negatieve haalbaarheidsfeedback meer dan twee derden (68.5%) van de studenten activeerde. Sommigen verhoogden hun inzet voor

hun studies en anderen overwogen om een andere opleiding te volgen. Van beide effecten wordt verwacht dat deze bevorderlijk zijn voor hun academisch traject.

Kort samengevat beschreven we de constructie van een valide instrument dat toelaat na te gaan welke specifieke opleidingen in overeenstemming zijn met de interesses van (toekomstige) studenten (hoofdstuk 3). In hoofdstuk 4 bevestigden we dat de basiskennis van wiskunde essentieel is voor academisch succes in het eerste jaar van het hoger onderwijs. In hoofdstuk 5 breidden we deze notie uit naar andere competenties door te tonen dat de assessment van basisvaardigheden toelaat studenten te identificeren die een zeer hoge kans hebben op falen in het eerste jaar hoger onderwijs. In hoofdstuk 6 toonden we dat het geven van feedback helpt om studenten met potentiële tekorten te waarschuwen en te activeren.

Deze hoofdstukken beschreven samen de basisprincipes en procedures van de ontwikkeling van SIMON, maar deze worden voortdurend toegepast en herhaald in grotere samples van studenten. Zo steunt de validering van de slaagkansen in SIMON (SIMON-C) ondertussen op gegevens van 22,008 studenten en bestaat de interessedatabank voor verdere validering van SIMON-I uit 13,535 respondenten. Jaarlijks start een nieuwe gegevensverzameling voor beide componenten. Dit garandeert continuïteit in de ontwikkeling en optimalisering van het instrument.

Implicaties en sterktes

De belangrijkste praktische implicatie van dit proefschrift is dat de resultaten werden vertaald in een instrument dat klaar is voor gebruik in een populatie van toekomstige studenten die een studiekeuze moeten maken. Naast dit instrument voor de initiële doelgroep wordt SIMON nu ook gebruikt bij studenten die reeds ingeschreven zijn in het hoger onderwijs.

SIMON voor studiekiezers: vraag het aan SIMON

SIMON is gratis beschikbaar voor iedereen via de website www.vraaghetaansimon.be. Gebruikers maken een persoonlijk profiel aan wat hen toelaat om steeds hun persoonlijke

gegevens te raadplegen en aan te vullen. Wanneer zij inloggen staat het hen vrij te kiezen met welke component zij van start gaan. Studiekiezers die onzeker zijn over wat zij willen studeren starten doorgaans met de interessevragenlijst. Gebruikers die weten wat hen interesseert, maar vragen hebben over hun kansen op succes gaan van start met SIMON-C. Deze vrije keuze impliceert tevens dat gebruikers SIMON steeds kunnen gebruiken en hergebruiken doorheen hun hele studiekeuzeproces.

Wanneer gebruikers de interessevragenlijst invullen krijgen zij hun persoonlijke interesseprofiel dat bestaat uit een grafische representatie van hun scores op de RIASEC-dimensies en op de academische schaal. Zij kunnen tevens bekijken welke opleidingen aansluiten bij hun interesses. De vragenlijst kan naar wens opnieuw ingevuld worden maar scores op de dimensies kunnen ook manueel aangepast worden waarna de lijst van overeenstemmende opleidingen verandert. Dit zorgt ervoor dat gebruikers opleidingen en hun kenmerken maximaal kunnen exploreren.

SIMON-C bestaat uit alle vaardigheidstesten die eerder werden gevalideerd. Wanneer een test is ingevuld krijgt de gebruiker een korte uitleg over de relevantie van de vaardigheid in het hoger onderwijs, samen met de persoonlijke score en de positie van de score ten opzichte van de referentiegroep van studiekiezers.

De resultaten van SIMON-I en SIMON-C worden geïntegreerd in het 'overzicht van opleidingen'. Deze pagina toont alle opgenomen opleidingen en de overeenstemming van deze opleidingen met de persoonlijke interesses en competenties. Deze overeenstemming wordt aangeduid met een kleurcode. Groen duidt op grote overeenstemming, rood op een kleine en oranje wijst op een gemiddelde *match*. Omdat de lijst met opleidingen omvangrijk is (155 programma's), kunnen gebruikers de opleidingen rangschikken op overeenstemming met interesses, met slaagkansen, op de aard van het programma (academisch of professioneel) of op alle voorgaande. Om gebruikers te helpen een selectie te maken van opleidingen kunnen zij

deze als 'favoriet' aanduiden en ervoor kiezen enkel deze opleidingen te zien. Een voorbeeld van dergelijke favorietenlijst met informatie over de overeenstemming met interesses en competenties wordt getoond in Figuur 1.

Fig. 1. Screenshot van vraag het aan SIMON waarbij de gebruiker aanduidde enkel de favorieten te willen zien. De eerste kolom toont de naam van de opleiding en de derde kolom toont of dit een academische dan wel een professionele opleiding betreft. De vierde kolom toont de overeenstemming met interesses en de laatste kolom toont de persoonlijke slaagkans in de opleiding.

Wanneer gebruikers op een opleiding klikken krijgen zij verdere details te zien zoals het interesseprofiel van succesvolle studenten in de betreffende opleiding en de instellingen die de opleiding organiseren. Wanneer zij op het logo van de instelling klikken worden zij doorverwezen naar de opleidingspagina waar ze meer informatie krijgen over de inhoud van de opleiding en over praktische details.

Gebruikers kunnen hun resultaten ook downloaden en printen.

SIMON is momenteel operationeel, maar mogelijkheden tot verdere uitbreiding (van testen, componenten, opleidingen…) werden geïncorporeerd. Het instrument wordt elk jaar ge-

updatet op basis van nieuwe data. De website werd gelanceerd in februari 2015 en werd in het eerste volledige schooljaar ('15-'16) gebruikt door 20,000 unieke studiekiezers. Ter vergelijking, 49,487 studenten schreven zich voor de eerste keer in in het hoger onderwijs in Vlaanderen in het academiejaar '15-'16 (Ministerie van Onderwijs en Vorming, 2016). Deze cijfers tonen dat studiekiezers nood hadden aan een valide instrument dat hen helpt bij de keuze van een opleiding hoger onderwijs.

Als resultaat van het succes van SIMON besliste de Vlaamse overheid een instrument te ontwikkelen met hetzelfde doel om studiekiezers te helpen bij hun keuzeproces: Columbus. Verschillende modules van SIMON werden opgenomen in de valideringsfase van Columbus. Dus, SIMON is niet enkel een gebruiksklaar instrument, maar was ook pionier in de ontwikkeling van een Vlaams, instellingsoverschrijdend studiekeuze-instrument.

SIMON voor studenten: SIMON zegt

Een tweede belangrijke praktische waarde van SIMON ligt in de toepassing bij reeds ingeschreven studenten. Sinds deze toepassing in 2015 werd geïntroduceerd kregen al 10,224 nieuw ingeschreven studenten aan de Universiteit Gent een persoonlijk feedbackrapport waarvan er 7,612 een persoonlijke slaagkans kregen voor de opleiding waarin ze waren ingeschreven. SIMON laat toe te identificeren welke studenten die een verhoogd risico hebben om te falen en het geeft hen een duidelijke waarschuwing. Bovendien geven de feedbackrapporten concreet en werkbaar advies over hoe studenten belangrijke vaardigheden kunnen bijschaven. Zoals aangetoond werd in hoofdstuk 6 missen deze rapporten hun doel niet. Zij stimuleren studenten actief bezig te zijn met hun studies of hun initiële keuze in vraag te stellen waardoor zij het studierendement in het hoger onderwijs kunnen verhogen.

Andere Sterktes en Implicaties

Een van de belangrijke voordelen van SIMON is de focus op lage slaagkansen. Gezien er echte basisvaardigheden worden getoetst ligt de focus op groeipotentieel en niet op de kennis

die al dan niet werd vergaard tijdens het secundair onderwijs. Gezien leerlingen met een lagere sociaal-economische status (SES) disproportioneel zijn ingeschreven in secundaire onderwijsrichtingen die minder goed voorbereiden op hoger onderwijs (Groenez et al., 2003), is de klemtoon op identificatie van basisvoorwaarden potentieel een sterke hulp in het promoten van sociale gelijkheid. Eén van de redenen waarom adolescenten met een lagere SES minder hoger onderwijs genieten is dat hun ouders een pessimistische kijk hebben op hun kansen op succes (Müller, 2014). SIMON kan deze barrière verhelpen door realistische informatie te geven en het zelfvertrouwen te versterken. Studiekiezers kunnen immers onverwacht een hoge slaagkans krijgen wat hen ertoe kan aanzetten een opleiding hoger onderwijs te volgen wanneer zij deze mogelijkheid eerst hadden uitgesloten. Op deze manier kan SIMON helpen om de sociale ongelijkheid in het hoger onderwijs weg te werken.

Het feit dat SIMON ook niet-cognitieve factoren in rekening brengt heeft eveneens belangrijke voordelen. Dit verhoogt niet alleen de accuraatheid van de voorspelling, maar het geeft (prospectieve) studenten ook een ruimer beeld van hun persoonlijke vaardigheden. Dit is belangrijk gezien niet-cognitieve factoren een gebrek aan cognitieve vaardigheden kunnen compenseren (Komarraju et al., 2013). De inclusie van niet-cognitieve variabelen schept ook mogelijkheden voor interventie. Gezien deze factoren belangrijk zijn voor succes in het hoger onderwijs, kan de ontwikkeling en organisatie van remediëringsactiviteiten om deze vaardigheden aan te scherpen het studierendement verhogen.

De differentiële predictieve validiteit van specifieke cognitieve en niet-cognitieve factoren in verschillende opleidingen heeft ook belangrijke implicaties en dit zowel voor begeleiding en advies als voor onderzoek. Zo dienen onderzoekers voorzichtig te zijn om bevindingen bij specifieke samples te generaliseren naar de volledige studentepopulatie. Met betrekking tot advies en begeleiding toonden we aan dat het mogelijk en wenselijk is om context-specifieke voorspellingen te doen, zelfs wanneer deze gebaseerd zijn op een uniforme

testbatterij. Dit is bijzonder nuttig voor studiekiezers die een enkel platform kunnen gebruiken om hun overeenstemming met vele opleidingen te evalueren. Dit verhoogt ook het exploreren van mogelijkheden gezien de gebruiker opleidingen te zien krijgt die hij/zij eerder nooit had overwogen.

Tot slot toonden we ook aan dat negatieve feedback studenten activeert om zich meer in te zetten voor hun studies of om een andere opleiding te overwegen, wat het studierendement verder kan verhogen.

Beperkingen en suggesties voor verder onderzoek

De belangrijkste toekomstige uitdaging ligt in een meer longitudinale opvolging van zowel interesses, competenties als feedback. Voor de interesses gaat dit over de longitudinale opvolging van leerlingen secundair onderwijs en nieuw ingeschreven studenten om de validiteit van SIMON-I voor studieresultaten te onderzoeken. Van studenten die zich inschrijven in een opleiding die bij hen past (volgens SIMON-I) wordt immers verwacht dat zij betere resultaten zullen neerzetten dan studenten waarvan SIMON de overeenstemming lager inschatte (Nye et al., 2012). Ook SIMON-C heeft baat bij een meer longitudinale aanpak. Tot op heden werd alleen succes in het eerste jaar hoger onderwijs voorspeld. Hoewel dit een sterke voorspeller is van verder succes (Allen, 1999; de Koning et al., 2012), moet onderzocht worden of de resultaten stand houden bij meer lange termijn criteria zoals het tijdig behalen van een diploma. Hetzelfde geldt voor de resultaten over het effect van feedback. Voorlopig is nog niet geweten of feedback een lange termijn effect heeft op assimilatie en accommodatie. Het is bijvoorbeeld mogelijk dat het krijgen van een lage slaagkans het effect van tegenvallende examenresultaten versterkt. Op die manier kan de feedback mogelijks een groter effect hebben op heroriëntering dan we tot op heden aannemen.

Een tweede belangrijke toekomstige ontwikkeling betreft de integratie van de interesse- en de competenties-componenten. Terwijl er bij de competenties gefocust wordt op de

voorspelling van academisch succes, beperkt de interessecomponent zich tot het weergeven van overeenstemmende opleidingen. Nochtans is de overeenstemming tussen interesses en opleidingen ook een voorspeller van academisch succes (Nye et al., 2012). Daarom is het aangewezen om te onderzoeken of de congruentie tussen interesses en opleidingen de accuraatheid van de voorspelling van opleidingsspecifieke slaagkansen kan verhogen. Dit zou studiekiezers ook een meer geïntegreerd beeld geven van de adequaatheid van specifieke opleidingen.

Een derde aanbeveling gaat over *fairness*. Hoewel SIMON geen selectie-instrument is en geen consequenties heeft voor toelating tot het hoger onderwijs, is het hoogst belangrijk dat het gegenereerde advies geen specifieke groepen bevoor- of benadeelt. Gezien het maatschappelijke doel van brede en open toegang tot opleidingen, is het begrijpen van het effect van SES op SIMON-resultaten bijzonder belangrijk (Sackett et al., 2012). In hoofdstuk 2 vonden we geen aanwijzingen dat SIMON bepaalde groepen bevoor- of benadeelt. Toch moet in de verdere ontwikkeling constant gemonitord worden of testresultaten geen ondergerepresenteerde groepen benadeelt.

Ten laatste, gezien feedback een wezenlijk onderdeel is bij het van bepalen van persoonlijke doelen, gaat de optimalisering van feedback hand in hand met een verhoogde kennis van het studiekeuzeproces in het algemeen en de individuele verschillen in heroriënteringsprocessen in het bijzonder. De processen en factoren die heroriëntering beïnvloeden zijn nog steeds onduidelijk (Creed & Blume, 2013; Heckhausen et al., 2010). Hoewel deze scriptie bijdroeg aan de kennis van het (her)oriënteringsproces blijft nog veel onduidelijk over welke factoren bepalen of iemand een onhaalbare opleiding blijft nastreven of zich ervan losmaakt (Boudrenghien et al., 2012; Carroll et al., 2009; Converse et al., 2010; Tomasik & Silbereisen, 2012). Toekomstig onderzoek moet uitwijzen hoe de aanvaarding van feedback kan gestimuleerd worden en hoe feedback heroriënteringsprocessen kan bevorderen.

Algemene conclusie

In deze doctoraatsscriptie werd de ontwikkeling en validering van SIMON beschreven, een instrument dat (toekomstige) studenten toelaat om na te gaan in welke mate specifieke opleidingen passen bij hun persoonlijke interesses en competenties.

De vier studies die werden beschreven behandelden de drie componenten van het instrument: de overeenstemming tussen de persoonlijke interesses en de opleidingen hoger onderwijs, de evaluatie van persoonlijks competenties in het licht van opleidingen, en het geven van activerende feedback.

Het onderzoek naar deze componenten droeg bij aan het begrip van academisch succes in het hoger onderwijs en van het studiekeuzeproces. Bovenal leidde dit onderzoek tot een valide instrument dat kan ingezet worden in de populatie van studiekiezers én bij ingeschreven studenten. SIMON ondersteunt zo de studiekeuze van jongeren en informeert studenten over de lacunes in hun kennis en vaardigheden. Op die manier kan het instrument bijdragen aan de verhoging van het studiesucces in het hoger onderwijs.

Referenties

Allen, D. (1999). Desire to finish college: an empirical link between motivation and persistence. *Research in Higher Education, 40*, 461-485.

Boudrenghien, G., Frenay, M., & Bourgeois, E. (2012). Unattainable educational goals: Disengagement, reengagement with alternative goals, and consequences for subjective well-being. *European Review of Applied Psychology-Revue Europeenne De Psychologie Appliquee, 62*(3), 147-159. doi:10.1016/j.erap.2012.04.002

Brandtstädter, J., & Rothermund, K. (2002). The life-course dynamics of goal pursuit and goal adjustment: A two-process framework. *Developmental Review, 22*(1), 117-150. doi:10.1006/drev.2001.0539

Brown, D. (2002). Introduction to Theories of Career Development and Choice: Origins, Evolution, and Current Efforts. In D. Brown (Ed.), *Career choice and development* (Fourth ed.). San Francisco: John Wiley and Sons, Inc.

Browne, M. W. (1992). Circumplex models for correlation matrices. *Psychometrika, 57*(4), 469-497. doi:10.1007/bf02294416

Carroll, P. J., Shepperd, J. A., & Arkin, R. M. (2009). Downward Self-Revision: Erasing Possible Selves. *Social Cognition, 27*(4), 550-578. doi:10.1521/soco.2009.27.4.550

Converse, P. D., Steinhauser, E., & Pathak, J. (2010). Individual differences in reactions to goal-performance discrepancies over time. *Personality and Individual Differences, 48*(2), 138-143. doi:10.1016/j.paid.2009.09.010

Credé, M., & Kuncel, N. (2008). Study habits, skills, and attitudes. The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science, 3*(6), 425-453. doi:10.1111/j.1745-6924.2008.00089.x

Creed, P. A., & Blume, K. (2013). Compromise, Well-Being, and Action Behaviors in Young Adults in Career Transition. *Journal of Career Assessment, 21*(1), 3-19. doi:10.1177/1069072712453830

de Koning, B., Loyens, S., Rikers, E., Smeets, G., & van der Molen, H. (2012). Generation Psy: Student characteristics and academic achievement in a three-year problem-based learning bachelor program. *Learning and Individual Differences, 22*(3), 313-323. doi:10.1016/j.lindif.2012.01.003

Dingemanse, S., Van Amstel, B., De Fruyt, F., & Wille, B. (2009). *Loopbaan Inzicht Vragenlijst - Herziene versie 2009*: Harcourt. Test publishers.

Duijnhouwer, H., Prins, F. J., & Stoking, K. M. (2012). Feedback providing improvement strategies and reflection on feedback use: Effects on students' writing motivation, process, and performance. *Learning and Instruction, 22*(3), 171-184. doi:10.1016/j.learninstruc.2011.10.003

Feldman, K., Smart, J., & Ethington, C. (1999). Major Field and Person-Environment Fit: Using Holland's theory to study change and stability of college students. *The Journal of Higher Education, 70*(6), 642-669. doi:10.2307/2649169

Hattie, J. (2009). Visibly learning from reports: the validity of score reports. *Online Educational Research Journal.* Retrieved from http://www.oerj.org/View?action=viewPDF&paper=6

Heckhausen, J., Wrosch, C., & Schulz, R. (2010). A motivational theory of life-span development. *Psychological review, 117*(1), 32. doi:10.1037/a0017668

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. Lutz, FL: Psychological Assessment Resources Inc.

Komarraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. *Learning and Individual Differences, 24*, 103-109. doi:10.1016/j.lindif.2012.12.007

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist, 50*(9), 741.

Ministerie van Onderwijs en Vorming. (2009). Studiesucces Generatiestudenten in 2007-2008. Retrieved from http://www.ond.vlaanderen.be/hogeronderwijs/werken/studentadmin/studentengegevens/Studiesucces%20generatiestudenten%20in%202007.pdf

Ministerie van Onderwijs en Vorming. (2016). Hoger Onderwijs in cijfers academiejaar 2015-2016. Retrieved from http://www.ond.vlaanderen.be/hogeronderwijs/werken/studentadmin/studentengegevens/HOIC_2015_2016.pdf

Müller, W. (2014). Educational inequality and social justice: Challenges for career guidance. *International Journal for Educational and Vocational Guidance, 14*(1), 21-33. doi:10.1007/s10775-014-9264-4

Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational Interests and Performance: A Quantitative Summary of Over 60 Years of Research. *Perspectives on Psychological Science, 7*(4), 384-403. doi:10.1177/1745691612449021

OECD. (2014). *Education at glance 2014: OECD Indicators.* OECD Publishing.

Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288. doi:10.1037/0033-2909.130.2.261

Tomasik, M. J., & Silbereisen, R. K. (2012). Beneficial effects of disengagement from futile struggles with occupational planning: A contextualist-motivational approach. *Developmental Psychology, 48*(6), 1785. doi:10.1037/a0026952

Tracey, T. J. G., & Rounds, J. B. (1993). Evaluating Holland's and Gati's vocational-interest models: A structural meta-analysis. *Psychological Bulletin, 113*(2), 229-246. doi:10.1037/0033-2909.113.2.229

Van Daal, T., Coertjens, L., Delvaux, E., Donche, V., & Van Petegem, P. (2013). *Klaar voor hoger onderwijs of arbeidsmarkt? Longitudinaal onderzoek bij laatstejaarsleerlingen secundair onderwijs.* Antwerpen - Apeldoorn: Garant.

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

Williams, K. J., Donovan, J. J., & Dodge, T. L. (2000). Self-regulation of performance: Goal establishment and goal revision processes in athletes. *Human Performance, 13*(2), 159-180. doi:10.1207/s15327043hup1302_3

Data Storage Fact Sheets

In compliance with the UGent standard for research accountability, transparency and reproducibility, the location of the datasets used in this dissertation are added below. For each of the empirical chapters (i.e., chapters 3 to 6) a separate Data Storage Fact Sheet is completed, detailing which data and analysis files are stored, where they are stored, who has access to the files and who can be contacted in order to request access to the files.

Data Storage Chapter 3

% Data Storage Fact Sheet

% Name/identifier study
% Author: Lot Fonteyne
% Date: 10/03/2017

1. Contact details
==

1a. Main researcher

- name: Lot Fonteyne
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Lot.Fonteyne@UGent.be

1b. Responsible Staff Member (ZAP)

- name: Filip De Fruyt
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Filip.DeFruyt@UGent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies
===
* Reference of the publication in which the datasets are reported:

Fonteyne, L., Wille, B., Duyck, W., & De Fruyt, F. (2016). Exploring vocational and academic fields of study: Development and validation of the Flemish SIMON Interest Inventory (SIMON-I). International Journal for Educational and Vocational Guidance(In Press). doi:10.1007/s10775-016-9327-9

* Which datasets in that publication does this sheet apply to?:

All datasets

3. Information about the files that have been stored
===

3a. Raw data
----

* Have the raw data been stored by the main researcher? [x] YES / [ ] NO
If NO, please justify:

* On which platform are the raw data stored?
  - [ ] researcher PC
  - [ ] research group file server
  - [x] other (specify): faculty database and file server

* Who has direct access to the raw data (i.e., without intervention of another person)?
  - [ ] main researcher and - [ ] responsible ZAP

- [x] all members of the research group

- [ ] all members of UGent

- [ ] other (specify): ...

3b. Other files

----

* Which other files have been stored?

 - [ ] file(s) describing the transition from raw data to reported results. Specify: ...

 - [x] file(s) containing processed data. Specify: spss file

 - [ ] file(s) containing analyses. Specify: ...

 - [ ] files(s) containing information about informed consent

 - [ ] a file specifying legal and ethical provisions

 - [ ] file(s) that describe the content of the stored files and how this content should be interpreted. Specify: ...

 - [ ] other files. Specify: ...

* On which platform are these other files stored?

 - [ ] individual PC

 - [ ] research group file server

 - [x] other: Project file server

* Who has direct access to these other files (i.e., without intervention of another person)?

 - [ ] main researcher

 - [ ] responsible ZAP

 - [x] all members of the research group

 - [ ] all members of UGent

 - [ ] other (specify): ...

4. Reproduction

```
============================================================
 * Have the results been reproduced independently?: [ ] YES / [x] NO


 * If yes, by whom (add if multiple):
   - name:
   - address:
   - affiliation:
 - e-mail:
```

Data Storage Chapter 4

```
% Data Storage Fact Sheet


% Name/identifier study
% Author: Lot Fonteyne
% Date: 10/03/2017



1. Contact details
============================================================


1a. Main researcher
-----------------------------------------------------------
- name: Lot Fonteyne
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Lot.Fonteyne@UGent.be


1b. Responsible Staff Member (ZAP)
-----------------------------------------------------------
- name: Filip De Fruyt
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Filip.DeFruyt@UGent.be
```

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

==

* Reference of the publication in which the datasets are reported:

Fonteyne, L., De Fruyt, F., Dewulf, N., Duyck, W., Erauw, K., Goeminne, K., . . . Rosseel, Y. (2014). Basic mathematics test predicts statistics achievement and overall first year academic success. European Journal of Psychology of Education, 1-24. doi:10.1007/s10212-014-0230-9

* Which datasets in that publication does this sheet apply to?:

All datasets

3. Information about the files that have been stored

==

3a. Raw data

----

* Have the raw data been stored by the main researcher? [x] YES / [ ] NO

If NO, please justify:

* On which platform are the raw data stored?

 - [ ] researcher PC

 - [ ] research group file server

 - [x] other (specify): project file server

* Who has direct access to the raw data (i.e., without intervention of another person)?

 - [ ] main researcher

- [ ] responsible ZAP

- [x] all members of the research group

- [ ] all members of UGent

- [ ] other (specify): ...

3b. Other files

----

* Which other files have been stored?

 - [ ] file(s) describing the transition from raw data to reported results. Specify: ...

 - [x] file(s) containing processed data. Specify: same file

 - [ ] file(s) containing analyses. Specify: ...

 - [ ] files(s) containing information about informed consent

 - [ ] a file specifying legal and ethical provisions

 - [ ] file(s) that describe the content of the stored files and how this content should be
interpreted. Specify: ...

 - [ ] other files. Specify: ...

 * On which platform are these other files stored?

 - [ ] individual PC

 - [ ] research group file server

 - [x] other: Project file server

 * Who has direct access to these other files (i.e., without intervention of another person)?

 - [ ] main researcher

 - [ ] responsible ZAP

 - [x] all members of the research group

 - [ ] all members of UGent

 - [ ] other (specify): ...

4. Reproduction

===

* Have the results been reproduced independently?: [ ] YES / [x] NO

* If yes, by whom (add if multiple):

  - name:

  - address:

  - affiliation:

- e-mail:

Data Storage Chapter 5

% Data Storage Fact Sheet

% Name/identifier study

% Author: Lot Fonteyne

% Date: 10/03/2017

1. Contact details

===

1a. Main researcher

- name: Lot Fonteyne

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: Lot.Fonteyne@UGent.be

1b. Responsible Staff Member (ZAP)

- name: Filip De Fruyt

- address: Henri Dunantlaan 2, 9000 Gent

- e-mail: Filip.DeFruyt@UGent.be

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

==

* Reference of the publication in which the datasets are reported:

Fonteyne, L., Duyck, W., & De Fruyt, F. (in press). *Program-specific prediction of academic achievement on the basis of cognitive and non-cognitive factors*. Learning and Individual Differences. doi: 10.1016/j.lindif.2017.05.003

* Which datasets in that publication does this sheet apply to?:

All datasets

3. Information about the files that have been stored

==

3a. Raw data

----

* Have the raw data been stored by the main researcher? [x] YES / [ ] NO
If NO, please justify:

* On which platform are the raw data stored?
 - [ ] researcher PC
 - [ ] research group file server
 - [x] other (specify): project file server

* Who has direct access to the raw data (i.e., without intervention of another person)?
 - [ ] main researcher
 - [ ] responsible ZAP

- [x] all members of the research group

- [ ] all members of UGent

- [ ] other (specify): ...

3b. Other files

----

* Which other files have been stored?

 - [ ] file(s) describing the transition from raw data to reported results. Specify: ...

 - [x] file(s) containing processed data. Specify: same file

 - [ ] file(s) containing analyses. Specify: ...

 - [ ] files(s) containing information about informed consent

 - [ ] a file specifying legal and ethical provisions

 - [ ] file(s) that describe the content of the stored files and how this content should be
interpreted. Specify: ...

 - [ ] other files. Specify: ...

* On which platform are these other files stored?

 - [ ] individual PC

 - [ ] research group file server

 - [x] other: Project file server

* Who has direct access to these other files (i.e., without intervention of another person)?

 - [ ] main researcher

 - [ ] responsible ZAP

 - [x] all members of the research group

 - [ ] all members of UGent

 - [ ] other (specify): ...

4. Reproduction

```
================================================================
* Have the results been reproduced independently?: [ ] YES / [x] NO


* If yes, by whom (add if multiple):
  - name:
  - address:
  - affiliation:
- e-mail:
```

Data Storage Chapter 6

```
% Data Storage Fact Sheet


% Name/identifier study
% Author: Lot Fonteyne
% Date: 10/03/2017


1. Contact details
================================================================


1a. Main researcher
-----------------------------------------------------------
- name: Lot Fonteyne
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Lot.Fonteyne@UGent.be


1b. Responsible Staff Member (ZAP)
-----------------------------------------------------------
- name: Filip De Fruyt
- address: Henri Dunantlaan 2, 9000 Gent
- e-mail: Filip.DeFruyt@UGent.be
```

If a response is not received when using the above contact details, please send an email to data.pp@ugent.be or contact Data Management, Faculty of Psychology and Educational Sciences, Henri Dunantlaan 2, 9000 Ghent, Belgium.

2. Information about the datasets to which this sheet applies

==

* Reference of the publication in which the datasets are reported:

Fonteyne, L., Eelbode, A., Lanszweert, I., Roels, E., Schelfhout, S., Duyck, W., & De Fruyt, F. (2017). Career goal engagement following negative feedback: influence of expectancy-value and perceived feedback accuracy. Manuscript submitted for publication.

* Which datasets in that publication does this sheet apply to?:

All datasets

3. Information about the files that have been stored

==

3a. Raw data

----

* Have the raw data been stored by the main researcher? [x] YES / [ ] NO

If NO, please justify:

* On which platform are the raw data stored?

  - [ ] researcher PC

  - [ ] research group file server

  - [x] other (specify): university database and project file server

* Who has direct access to the raw data (i.e., without intervention of another person)?

  - [ ] main researcher and - [ ] responsible ZAP

  - [x] all members of the research group

- [ ] all members of UGent

  - [ ] other (specify): ...

3b. Other files

----

* Which other files have been stored?

 - [ ] file(s) describing the transition from raw data to reported results. Specify: ...

 - [x] file(s) containing processed data. Specify: spss file

 - [ ] file(s) containing analyses. Specify: ...

 - [ ] files(s) containing information about informed consent

 - [ ] a file specifying legal and ethical provisions

 - [ ] file(s) that describe the content of the stored files and how this content should be
interpreted. Specify: ...

 - [ ] other files. Specify: ...

 * On which platform are these other files stored?

 - [ ] individual PC

 - [ ] research group file server

 - [x] other: Project file server

 * Who has direct access to these other files (i.e., without intervention of another person)?

 - [ ] main researcher

 - [ ] responsible ZAP

 - [x] all members of the research group

 - [ ] all members of UGent

 - [ ] other (specify): ...

4. Reproduction

===

* Have the results been reproduced independently?: [ ] YES / [x] NO

* If yes, by whom (add if multiple):

  - name:

  - address:

  - affiliation:

 - e-mail: